

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 06:43:25 ; Search time 27 Seconds
(without alignments)
1233.000 Million cell updates/sec

Title: US-10-070-277-2

Perfect score: 1812

Sequence: 1 MEISTQPDMDHLHRLDGV.....SGDIIMFAGEMDMLPAVL 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1467	81.0	377	2	T05124
2	1056.5	58.3	342	2	S76410
3	1039.5	57.4	350	2	E82198
4	983	54.2	344	2	G81934
5	982	54.2	344	2	H81171
6	975	53.8	348	2	T10453
7	971	53.6	348	1	DEECOO
8	970	53.5	348	2	DB5668
9	970	53.5	348	2	H90808
10	956	52.8	348	2	AF0193
11	951	52.5	393	2	AF2625
12	951	52.5	393	2	E97407
13	949	52.4	393	2	AD0638
14	946	52.2	348	1	DEEBOT
15	859.5	47.4	350	2	G84968
16	569	31.4	335	2	B81444
17	566.5	31.3	337	2	T37731
18	541	29.9	339	2	E64592
19	527	29.1	339	2	A71922
20	481	26.5	364	1	DEBYO
21	466.5	25.7	391	1	DEUSO
22	146	8.1	423	2	A64486
23	140.5	7.8	422	2	C70370
24	139	7.7	428	2	B83967
25	131.5	7.3	426	2	A81304
26	131	7.2	430	2	T30519
27	130.5	7.2	430	2	T45151
28	129.5	7.1	426	2	A81676
29	129.5	7.1	434	2	A69017

ALIGNMENTS

RESULT 1

T05124 dihydroorotase (EC 3.5.2.3) - Arabidopsis thaliana

N;Alternate names: protein F7H19.110

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T05124

R;Bevan, M.; Peters, S.A.; Van Staeren, M.; Dirke, W.; Stiekema, W.; Bancroft, I.; Me

submitted to the Protein Sequence Database, July 1998

A;Reference number: Z15399

A;Accession: T05124

A;Molecule type: DNA

A;Residues: 1-377 <BEV>

A;Cross-references: UNIPROT:O04904; EMBL:AL031018

A;Experimental source: cultivar Columbia; BAC clone F7H19

C;Genetics:

A;Map position: 4

A;Intons: 10/3; 60/2; 121/1; 166/3; 222/1; 261/1; 314/3

A;Note: F7H19.110

C;Function:

A;Description: catalyzes conversion of N-carbamoyl-L-aspartate into dihydroorotase

C;Keywords: hydrolase; pyrimidine nucleotide biosynthesis; zinc

Query Match 81.0%; Score 1467; DB 2; Length 377;

Best Local Similarity 78.4%; Pred. No. 1.5e-113;

Matches 269; Conservative 41; Mismatches 33; Indels 0; Gaps 0;

QY	1	MEISTQPDMDHLHRLDGV	1	MEISTQPDMDHLHRLDGV	1	MEISTQPDMDHLHRLDGV
DB	33	MEISTQPDMDHLHRLDGV	33	MEISTQPDMDHLHRLDGV	33	MEISTQPDMDHLHRLDGV
QY	61	SLPVDSPFNLMTLYLTDTTSPMEIKLARSQVFGVGLYPAGATTSGDVTDFGCL	60	SLPVDSPFNLMTLYLTDTTSPMEIKLARSQVFGVGLYPAGATTSGDVTDFGCL	60	SLPVDSPFNLMTLYLTDTTSPMEIKLARSQVFGVGLYPAGATTSGDVTDFGCL
DB	93	ALPESSEFDPMLTYLTDTTSPMEIKLARSQVFGVGLYPAGATTSGDVTDFGCL	152	ALPESSEFDPMLTYLTDTTSPMEIKLARSQVFGVGLYPAGATTSGDVTDFGCL	152	ALPESSEFDPMLTYLTDTTSPMEIKLARSQVFGVGLYPAGATTSGDVTDFGCL
QY	121	PVLQEMVHNHMPPLVHGEVNTPEVMDREKVFLETVLRLPVQKFPOLKVMMEHTTIDA	180	PVLQEMVHNHMPPLVHGEVNTPEVMDREKVFLETVLRLPVQKFPOLKVMMEHTTIDA	180	PVLQEMVHNHMPPLVHGEVNTPEVMDREKVFLETVLRLPVQKFPOLKVMMEHTTIDA
DB	153	PVLEEMVKQMPPLVHGEVNTPEVMDREKVFLETVLRLPVQKFPOLKVMMEHTTIDA	212	PVLEEMVKQMPPLVHGEVNTPEVMDREKVFLETVLRLPVQKFPOLKVMMEHTTIDA	212	PVLEEMVKQMPPLVHGEVNTPEVMDREKVFLETVLRLPVQKFPOLKVMMEHTTIDA
QY	181	VKPYESTGEGVAAATTPQHLVLRNSLPQGGLOPHNYCVPVLRKREIHRRLVAATVSGS	240	VKPYESTGEGVAAATTPQHLVLRNSLPQGGLOPHNYCVPVLRKREIHRRLVAATVSGS	240	VKPYESTGEGVAAATTPQHLVLRNSLPQGGLOPHNYCVPVLRKREIHRRLVAATVSGS
DB	213	VNFVESCSEGSVGTATVPQHLVLRNSLPQGGLOPHNYCVPVLRKREIHRRLVAATVSGS	272	VNFVESCSEGSVGTATVPQHLVLRNSLPQGGLOPHNYCVPVLRKREIHRRLVAATVSGS	272	VNFVESCSEGSVGTATVPQHLVLRNSLPQGGLOPHNYCVPVLRKREIHRRLVAATVSGS
QY	241	KRFLGTDSAPHRDRKESGCGAGIYNAPVALSVAKVEKENALDKLEAFSFGPDF	300	KRFLGTDSAPHRDRKESGCGAGIYNAPVALSVAKVEKENALDKLEAFSFGPDF	300	KRFLGTDSAPHRDRKESGCGAGIYNAPVALSVAKVEKENALDKLEAFSFGPDF
DB	273	KRFLGTDSAPHRDRKESGCGAGIYNAPVALSVAKVEKENALDKLEAFSFGPDF	332	KRFLGTDSAPHRDRKESGCGAGIYNAPVALSVAKVEKENALDKLEAFSFGPDF	332	KRFLGTDSAPHRDRKESGCGAGIYNAPVALSVAKVEKENALDKLEAFSFGPDF
QY	301	YGLPRNSKIKLSTPKVPSFYSASGDIIMFAGEMDMLP	343	YGLPRNSKIKLSTPKVPSFYSASGDIIMFAGEMDMLP	343	YGLPRNSKIKLSTPKVPSFYSASGDIIMFAGEMDMLP
DB	333	YGLPRNSKIKLSTPKVPSFYSASGDIIMFAGEMDMLP	375	YGLPRNSKIKLSTPKVPSFYSASGDIIMFAGEMDMLP	375	YGLPRNSKIKLSTPKVPSFYSASGDIIMFAGEMDMLP

RESULT 2

S76410

A:Reference: EC 3.5.2.3 - Synecocystis sp. (strain FCC 6803)

A:Accession: S76410

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-342 <KAN>

A:Cross-references: UNIPROT:P74438; EMBL:D90915; GB:AB001339; NID:g1653604; PTDN:BA1853

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Function:

A:Description: catalyzes conversion of N-carbamoyl-L-aspartate into dihydroorotate

A:Pathway: pyrimidine nucleotide biosynthesis

A:Note: zinc required for catalytic activity

C:Superfamily: dihydroorotase

C:Keywords: hydrolase; pyrimidine nucleotide biosynthesis; zinc

F:3,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

F:13,15/Binding site: zinc

Query Match

Best Local Similarity

Matches 200; Conservative

Score 1039.5; DB 2; Length 350;

Pred. No. 1.4e-73; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Mismatches 45; Mismatches 91; Indels 3; Gaps 3;

Query Match

Best Local Similarity

Matches 192; Conservative

Score 983; DB 2; Length 344;

Pred. No. 1.4e-73; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Mismatches 88; Mismatches 88; Indels 6; Gaps 5;

Db 122 LEBMAKQGLFLVHGVTPDEIDIFREAAFIKRWKPVLAQVYNLKVPEHITTEAAR 181
Qy 183 FVESCETGEVAAATVTPQHLVLNNSLFOGGLQPHNYCLPVLKREIHEALVSAVTSGRK 241
Db 182 LVLAEADN-VAAATVTPQHLVLNNSLFOGGLQPHNYCLPVLKREIHEALVSAVTSGRK 239
Qy 242 -RFLGTDSAPHDRRRKESCGCAGIYNAPVALSVYAKVFEKNALDKLEAFISFNGPDF 300
Db 240 HKFELGTDSAPHDRRRKESCGCAGIYNAPVALSVYAKVFEKNALDKLEAFISFNGPDF 299
Qy 301 YGLPRNNSKIKLSKTPMKVPESSFSYASGDIIPMFAGEMLDW 341
Db 300 YGIPENTDTITLVKQSQVTPASIPYGDDELVPMAAGSEIGW 340

RESULT 5

H81171
dihydroorotase NMB0682 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C/Accession: H81171
Ritterling, H.; Sauderer, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Hatt, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; et al., H.; Qin, H.; Vanachavan, J.; Gill, U.; Scarlato, V.; Maignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; et al. Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: H81171
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-344 <STO>
A:Cross-references: UNIPROT:Q9K0D1; GB:AE002422; GB:AE002098; NID:G7225898; PIDN:AAE4110
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0682
C:Superfamily: dihydroorotase

Query Match

54.2%; Score 982; DB 2; Length 344;
Best Local Similarity 56.3%; Pred. No. 1.7e-73;

Matches 192; Conserva 55; Mismatches 88; Indels 6; Gaps 5;

Qy 3 LSTTPDDMHLNRDGVLLKAVVSHSAHFGRATVMPNLPKPTTTTAAVAAYREAILKSL 62
Db 4 LTIIRPDMDHILNRDGVLLKAVVSHSAHFGRATVMPNLPKPTTTTAAVAAYREAILKSL 63
Qy 63 PVSDENPLMTLYLTDTTSPMEIKLARESOVFGVKLYPAGATTNSQDGVTDLEFGKCLPV 122
Db 64 PEGSAFEPMLTLYLTDTTSPMEIKLARESOVFGVKLYPAGATTNSQDGVTDLEFGKCLPV 121
Qy 123 LQEWVHNMLPLVHGEVTNPEVDVDFREKVFIEITVLRPLVQKFPOLKVMHEVTTIDAVK 182
Db 122 LEBMAKQGLFLVHGVTPDEIDIFREAAFIKRWKPVLAQVYNLKVPEHITTEAAR 181
Qy 183 FVESCETGEVAAATVTPQHLVLNNSLFOGGLQPHNYCLPVLKREIHEALVSAVTSGRK 241
Db 182 LVLAEADN-VAAATVTPQHLVLNNSLFOGGLQPHNYCLPVLKREIHEALVSAVTSGRK 239
Qy 242 -RFLGTDSAPHDRRRKESCGCAGIYNAPVALSVYAKVFEKNALDKLEAFISFNGPDF 300
Db 240 HKFELGTDSAPHDRRRKESCGCAGIYNAPVALSVYAKVFEKNALDKLEAFISFNGPDF 299
Qy 301 YGLPRNNSKIKLSKTPMKVPESSFSYASGDIIPMFAGEMLDW 341
Db 300 YGIPENTDTITLVKQSQVTPASIPYGDDELVPMAAGSEIGW 340

RESULT 6

dhhydroorotase (EC 3.5.2.3) - Pseudomonas aeruginosa
C/Species: Pseudomonas aeruginosa
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T10453; E83204

R.Brichta, D.M.; Brown, T.M.; Houghton, J.E.; O'Donovan, G.A.
submitted to the EMBL Data Library, November 1998
A:Description: Pyrc (dihydroorotase) sequence from P. aeruginosa.
A:Reference number: Z17024
A:Accession: T10453
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-348 <BRI>
A:Cross-references: UNIPROT:P72170; EMBL:U73505; NID:G3868711; PID:G3868712
R.Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; Lim, J.; Lory, S.; Olson, M.V. Nature 405, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen. A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83204
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <STO>
A:Cross-references: GB:AE004773; GB:AE004091; NID:G9949672; PIDN:AA06915.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: pyrc; PA3527
C:Function:
A:Description: catalyzes conversion of N-carbamoyl-L-aspartate into dihydroorotase
A:Pathway: pyrimidine nucleotide biosynthesis
C:Superfamily: dihydroorotase
C:Keywords: hydrolase; pyrimidine nucleotide biosynthesis; zinc

Query Match

53.8%; Score 975; DB 2; Length 348;
Best Local Similarity 55.5%; Pred. No. 6.4e-73;

Matches 188; Conserva 43; Mismatches 106; Indels 2; Gaps 2;

Qy 3 LSTTPDDMHLNRDGVLLKAVVSHSAHFGRATVMPNLPKPTTTTAAVAAYREAILKSL 62
Db 5 LTIIRPDMDHILNRDGVLLKAVVSHSAHFGRATVMPNLPKPTTTTAAVAAYREAILKSL 64
Qy 63 PVSDENPLMTLYLTDTTSPMEIKLARESOVFGVKLYPAGATTNSQDGVTDLEFGKCLPV 122
Db 65 PAASRFEPLMTLYLTDTTSPMEIKLARESOVFGVKLYPAGATTNSQDGVTDLEFGKCLPV 123
Qy 123 LQEWVHNMLPLVHGEVTNPEVDVDFREKVFIEITVLRPLVQKFPOLKVMHEVTTIDAVK 182
Db 124 LEBMAKQGLFLVHGVTPDEIDIFREAAFIKRWKPVLAQVYNLKVPEHITTEAAR 183
Qy 183 FVESCETGEVAAATVTPQHLVLNNSLFOGGLQPHNYCLPVLKREIHEALVSAVTSGRK 242
Db 184 FVREAPAN-VGATITTAHLLYNRNHMLVGIRPHFCLPLIKRTHQDALDAVSGNPK 242
Qy 243 FFLGTDSAPHDRRRKESCGCAGIYNAPVALSVYAKVFEKNALDKLEAFISFNGPDFYG 302
Db 243 FFLGTDSAPHDRRRKESCGCAGIYNAPVALSVYAKVFEKNALDKLEAFISFNGPDFYG 302
Qy 303 LPRNNSKIKLSKTPMKVPESSFSYASGDIIPMFAGEMLDW 341
Db 303 LPRNNSKIKLSKTPMKVPESSFSYASGDIIPMFAGEMLDW 341

RESULT 7

dhhydroorotase (EC 3.5.2.3) - Escherichia coli (strain K-12)
N/Alternate names: carbamoylasepartic dehydrase
C/Species: Escherichia coli
C/Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C/Accession: A25008; A27084; C64849
R.Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; Lim, J.; Lory, S.; Olson, M.V. Nature 405, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen. A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: A25008
A:Molecule type: DNA
A:Residues: 1-348 <BAC>
A:Cross-references: UNIPROT:P05020; EMBL:X04469; NID:G42605; PIDN:CAA28157.1; PID:G4260

A:Experimental source: strain K-12
 R:Wilson, H.R.; Chan, P.T.; Turbough Jr., C.L.
 J. Bacteriol. 169, 3051-3058, 1987
 A:Title: Nucleotide sequence and expression of the *pyrC* gene of *Escherichia coli* K-12.
 A:Reference number: A91836; MUID:87250268; PMID:2885307
 A:Accession: A27084
 A:Molecule type: mRNA
 A:Residues: 1-348 <WIL>
 A:Cross-references: EMBL:M6752; NID:g147472; PIDN:AAA24482.1; PID:g147473
 A:Experimental source: strain K-12
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: C64849
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-348 <BLAT>
 A:Cross-references: GB:AE00207; GB:U00096; NID:g1787293; PIDN:AAC74146.1; PID:g1787301
 A:Experimental source: strain K-12, substrain MG1655
 C:Comment: The *de novo* synthesis of UMP, the precursor of all pyrimidine nucleotides, is
 ressed primarily by a cytidine nucleotide.
 C:Genetics:
 A:Gene: *pyrC*
 A:Map position: 23 min
 A:Function:
 A:Description: catalyzes conversion of N-carbamoyl-L-aspartate into dihydroorotate
 A:Pathway: pyrimidine nucleotide biosynthesis
 A>Note: zinc required for catalytic activity
 C:Superfamily: dihydroorotase
 C:Keywords: homodimer; hydrolase; pyrimidine nucleotide biosynthesis; zinc
 F:17.19/Binding site: zinc (His) #status predicted

Query Match 53.6%; Score 971; DB 1; Length 348;
 Best Local Similarity 55.8%; Pred. No. 1.4e-72;
 Matches 189; Conservative 46; Mismatches 102; Indels 2; Gaps 2;

QY 3 LSTGPDMMHLRGGDVLRKAVSHSAHFGRAIVMNLKPPITTTAAVAAYREAILKSL 62
 DB 8 LKIRRPDDMMHLRGGDVLRKAVSHSAHFGRAIVMNLKPPITTTAAVAAYREAILKSL 62
 QY 63 PVDSDPNLMTLYLDTTSPMEIKLARSQVFGVKLYPAGATTNSODGVTDFGKCLPV 122
 DB 68 PAGHDFPLMTCYLTDLSLDPELERNEGVFTAAKLYPAAATTNSHGVTSV-DAIMPV 126
 QY 123 LOEWVHNMPPLVHGEVTNPEDVMEFDEKVFIEVTLRPLVOKFPOLKVMHEVTTIDAVK 182
 DB 127 LERMEKIGMPPLVHGEVTNPEDVMEFDEKVFIEVTLRPLVOKFPOLKVMHEVTTIDAVK 182
 QY 183 FVESCTEGFVAATVTPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSQSKR 242
 DB 187 YVRDGNR-LAATITTPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSQSKR 242
 QY 243 FFLGDSAPHRDRRKECCGACAGIYVAPVALSVYAKVFENALDKLEAFTSFGNDPFG 302
 DB 246 VFLGDSAPHRDRRKECCGACAGIYVAPVALSVYAKVFENALDKLEAFTSFGNDPFG 302
 QY 303 LPRNNSKIKLSKTPMWKPPESFSYASGDIIPMPAGEMLDW 341
 DB 306 LPVNDTFIELVREHQVVASIALTDITLVPLAAGETVRW 344

RESULT 8
 DB5668
 dihydro-orotase [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)
 C:Species: *Escherichia coli*
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: DB5668
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: DB5668
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-348 <STO>
 A:Cross-references: UNIPROT:Q8X8N8; GB:AE005174; NID:g12514597; PIDN:AAG55808.1; GSPDB
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: *pyrC*
 C:Superfamily: dihydroorotase

Query Match 53.5%; Score 970; DB 2; Length 348;
 Best Local Similarity 55.8%; Pred. No. 1.7e-72;
 Matches 189; Conservative 46; Mismatches 102; Indels 2; Gaps 2;

QY 3 LSTGPDMMHLRGGDVLRKAVSHSAHFGRAIVMNLKPPITTTAAVAAYREAILKSL 62
 DB 8 LKIRRPDDMMHLRGGDVLRKAVSHSAHFGRAIVMNLKPPITTTAAVAAYREAILKSL 62
 QY 63 PVDSDPNLMTLYLDTTSPMEIKLARSQVFGVKLYPAGATTNSODGVTDFGKCLPV 122
 DB 68 PAGHDFPLMTCYLTDLSLDPELERNEGVFTAAKLYPAAATTNSHGVTSV-DAIMPV 126
 QY 123 LOEWVHNMPPLVHGEVTNPEDVMEFDEKVFIEVTLRPLVOKFPOLKVMHEVTTIDAVK 182
 DB 127 LERMEKIGMPPLVHGEVTNPEDVMEFDEKVFIEVTLRPLVOKFPOLKVMHEVTTIDAVK 182
 QY 183 FVESCTEGFVAATVTPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSQSKR 242
 DB 187 YVRDGNR-LAATITTPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSQSKR 242
 QY 243 FFLGDSAPHRDRRKECCGACAGIYVAPVALSVYAKVFENALDKLEAFTSFGNDPFG 302
 DB 246 VFLGDSAPHRDRRKECCGACAGIYVAPVALSVYAKVFENALDKLEAFTSFGNDPFG 302
 QY 303 LPRNNSKIKLSKTPMWKPPESFSYASGDIIPMPAGEMLDW 341
 DB 306 LPVNDTFIELVREHQVVASIALTDITLVPLAAGETVRW 344

RESULT 9
 H90808
 dihydro-orotase [imported] - *Escherichia coli* (strain O157:H7, substrain R1MD 0509952)
 C:Species: *Escherichia coli*
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: H90808
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Iehi, K.; Yokoyama, K.; Han, C.G
 gasakara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: H90808
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-348 <HAY>
 A:Cross-references: UNIPROT:Q8X8N8; GB:BA000007; PIDN:BA834863.1; PID:g13360904; GSPDB:C
 A:Experimental source: strain O157:H7, substrain R1MD 0509952
 C:Genetics:
 A:Gene: *ECs1440*
 C:Superfamily: dihydroorotase

Query Match 53.5%; Score 970; DB 2; Length 348;
 Best Local Similarity 55.8%; Pred. No. 1.7e-72;
 Matches 189; Conservative 46; Mismatches 102; Indels 2; Gaps 2;

QY 3 LSTGPDMMHLRGGDVLRKAVSHSAHFGRAIVMNLKPPITTTAAVAAYREAILKSL 62
 DB 8 LKIRRPDDMMHLRGGDVLRKAVSHSAHFGRAIVMNLKPPITTTAAVAAYREAILKSL 62
 QY 63 PVDSDPNLMTLYLDTTSPMEIKLARSQVFGVKLYPAGATTNSODGVTDFGKCLPV 122
 DB 68 PAGHDFPLMTCYLTDLSLDPELERNEGVFTAAKLYPAAATTNSHGVTSV-DAIMPV 126
 QY 123 LOEWVHNMPPLVHGEVTNPEDVMEFDEKVFIEVTLRPLVOKFPOLKVMHEVTTIDAVK 182
 DB 127 LERMEKIGMPPLVHGEVTNPEDVMEFDEKVFIEVTLRPLVOKFPOLKVMHEVTTIDAVK 182
 QY 183 FVESCTEGFVAATVTPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSQSKR 242
 DB 187 YVRDGNR-LAATITTPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSQSKR 242
 QY 243 FFLGDSAPHRDRRKECCGACAGIYVAPVALSVYAKVFENALDKLEAFTSFGNDPFG 302
 DB 246 VFLGDSAPHRDRRKECCGACAGIYVAPVALSVYAKVFENALDKLEAFTSFGNDPFG 302
 QY 303 LPRNNSKIKLSKTPMWKPPESFSYASGDIIPMPAGEMLDW 341
 DB 306 LPVNDTFIELVREHQVVASIALTDITLVPLAAGETVRW 344

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 01:28:38 ; Search time 5666 Seconds
(without alignments)
10869.499 Million cell updates/sec

Title: US-10-070-277-1

Perfect score: 1271

Sequence: 1 ttgcataaatgagctctca.....tgaagattcactgataaaaa 1271

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:1:
1: gb_Da:1:
2: gb_Htg:1:
3: gb_In:1:
4: gb_Om:1:
5: gb_Ov:1:
6: gb_Pat:1:
7: gb_Pn:1:
8: gb_Pl:1:
9: gb_Pr:1:
10: gb_Ro:1:
11: gb_Sts:1:
12: gb_Sy:1:
13: gb_Un:1:
14: gb_Vi:1:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1271	100.0	1271	6	AX088927 Sequence
2	1271	100.0	1271	6	AX093580 Sequence
3	646	50.8	1743	6	AX470409 Sequence
4	643.2	50.6	1038	8	BVU251897 Beta vulg
5	587	46.2	1362	8	ATPAF00146 Arabidops
6	549.2	43.2	1750	8	AK122010 Oryza sat
7	345.6	27.2	130001	1	D90915 Synecocyst
8	294.2	23.1	306050	1	BX321858 Nitrosomo
9	290.6	22.9	303450	1	AP005085 Chromobac
10	282.6	22.2	208524	1	AR375107 Sequence
11	281	22.1	1071	6	AE004420 Vibrio ch
12	281	22.1	11619	1	AE004420 Prochloro
13	280.4	22.1	349082	1	BX572091 Yersinia
14	277.2	21.8	10421	1	AE013777 Yersinia
15	277.2	21.8	110000	1	BX936398 Continuation (30 o
16	277.2	21.8	193050	1	AJ141449 Yersinia
17	277.2	21.8	291326	1	AE017135 Yersinia
18	276.4	21.7	348672	1	CR378670 Photobact
19	275.2	21.7	190050	1	AL646059 Ralstonia

20	274.6	21.6	110000	1	AE017282_22
21	273	21.5	253750	1	AP005345 Vibri
22	272.8	21.5	182624	6	AX770901 Sequence
23	272.8	21.5	348613	1	BX571865 Photobac
24	269.8	21.2	336182	1	AE016813 Vibri
25	267.4	21.0	1041	6	AR318865 Sequence
26	265.8	20.9	300355	1	AE017162 Prochloro
27	265.2	20.9	110000	1	BX571965 Continuation (35 o
28	265.2	20.9	110000	1	CP000010_25 Continuation (26 o
29	264.2	20.8	110000	6	AR409405_3 Continuation (4 of
30	264.2	20.8	110000	6	BD061520_3 Continuation (4 of
31	264.2	20.8	293181	1	AP001119 Buchera
32	263.6	20.7	349213	1	BX569693 Synecoco
33	262.8	20.7	110000	1	BX950851_20 Continuation (21 o
34	261.4	20.6	349323	1	BX572096 Prochloro
35	260.6	20.5	10909	1	AE007977 Agrobacte
36	260.6	20.5	13047	1	AE009010 Agrobacte
37	257	20.2	10986	1	AE015804 Shewanell
38	256.2	20.2	11088	1	AE014109 Buchera
39	254.8	20.0	300490	1	AE016965 Coxiiella
40	245.6	19.3	1629	1	ECOPYRCA X0469 E. coli
41	245.6	19.3	2046	1	ECOPYRCA X0469 E. coli
42	245.6	19.3	12029	1	AE005316 Escherichia
43	245.6	19.3	19201	1	D90742 Escherichia
44	245.6	19.3	19418	1	D90743 Escherichia
45	245.6	19.3	110000	1	U00096_11 Continuation (12 o

ALIGNMENTS

RESULT 1
AX088927 1271 bp DNA linear PAT 17-MAR-2001
LOCUS Sequence 1 from Patent WO0114569.
DEFINITION AX088927
ACCESSION AX088927
VERSION AX088927.1 GI:13397685
KEYWORDS
SOURCE
ORGANISM
Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE
AUTHORS
1. Ehrhardt, T., Stitt, M., Geigenberger, P. L., Loeff, I., Ziemer, R. and Schreider, M.
TITLE
Increasing the polysaccharide content in plants
JOURNAL
Patent: WO 0114569-A1 01-MAR-2001;
BASIS AKTENGESCHLACHT (DE)

FEATURES
source
Location/Qualifiers
1. 1271
/organism="Solanum tuberosum"
/mol_type="unassigned DNA"
/db_xref="taxon:4113"
9..1049
/note="unassigned protein product"
/codon_start=1
/protein_id="CAC34708.1"
/db_xref="GI:13397686"
/translation="MELSLTPQDPDMLHLRDGVLKAVSHSAHFGRAIYVNLKPP
ITTTAAVAAYREAILKIPVSDENPLMTLYLPTTSPMEIKLAREQVGVGLYPA
GATNSQDGVTDLPKGLPYLOKVEHNMPLVHGEVTNEVDNDEKIVEIYLRP
LVQKFPOLKVEHVTITDAVKPVESTEGSVATVYPOHVINRSLFQGGLOPHNY
CLPYLKEIHREAVSVAIVSGSKRFPGTDSAPDRRREKSCCGACAGIYNAPVALSVY
AKVFEKENALDKLEAFISFNGPDYGLPRNNSKIKIKTKTPKVPESFVASGDIIPMF
AGEMLDWLPAPL"

CDS

ORIGIN
Query Match 100.0%; Score 1271; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 4.8e-313; Indels 0; Gaps 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TTGCAAAAATGAGCTCTCATTCACACAACTGATGATTCATCTTCGCTGATG 60

Db 1 TTGCAAAATGAGAGCTCTCAATACACACCTGAGATGGCATCTTCACTCCCGTAGTG 60
Qy 61 GTATGTTCTTAAAGCAGTTGTCTCTCAGAGTGCAATCATCTTGGAGGGCAATATGCA 120
Db 61 GTATGTTCTTAAAGCAGTTGTCTCTCAGAGTGCAATCATCTTGGAGGGCAATATGCA 120
Qy 121 TGGCAAAATTTGAAGCCTCTATCACTACACCTGCTGCTGTAGCAATACCGGAGGCGA 180
Db 121 TGGCAAAATTTGAAGCCTCTATCACTACACCTGCTGCTGTAGCAATACCGGAGGCGA 180
Qy 181 TATGAAATCTTAACTGCTGTAGTAGATTTCAACCTCTTATAGCACTTATTTAGAG 240
Db 181 TATGAAATCTTAACTGCTGTAGTAGATTTCAACCTCTTATAGCACTTATTTAGAG 240
Qy 241 ATACAAACGAGCTCTATGGAATCAATACGAAAGAGAGCGAGTGTATTTGGGGTGA 300
Db 241 ATACAAACGAGCTCTATGGAATCAATACGAAAGAGAGCGAGTGTATTTGGGGTGA 300
Qy 301 AGTTGACCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 360
Db 301 AGTTGACCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 360
Qy 361 AGTTGACCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 420
Db 361 AGTTGACCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 420
Qy 421 AGTTGACCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 480
Db 421 AGTTGACCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 480
Qy 481 TAAAGACCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 540
Db 481 TAAAGACCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 540
Qy 541 TTGATGCTGCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 600
Db 541 TTGATGCTGCTGCTGCTGCTGCAAGCAATTTCTCAAGATGAGTGAATCTTCTTGGGA 600
Qy 601 CACAAACATCTTGTGTTGAAGCAAGCAATTTCTCTTCCAGGGGCTTACAAACGCAATAT 660
Db 601 CACAAACATCTTGTGTTGAAGCAAGCAATTTCTCTTCCAGGGGCTTACAAACGCAATAT 660
Qy 661 ACTGCTTCCAGTCTCTGAAAGAGAGATCCACAGGAGGCACTTGTGCTGCTGTAACA 720
Db 661 ACTGCTTCCAGTCTCTGAAAGAGAGATCCACAGGAGGCACTTGTGCTGCTGTAACA 720
Qy 721 GTGGAAGTAAAGATTTTCTTGGGAGTGAATGCTCTCTCATGATAGACGAAGAAAG 780
Db 721 GTGGAAGTAAAGATTTTCTTGGGAGTGAATGCTCTCTCATGATAGACGAAGAAAG 780
Qy 781 AGTGTCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 AGTGTCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy 841 AGTGTCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 AGTGTCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 901 CAGATTTTATGAGGCTCTGAGAACTCAAGATTAAGTTAGTGAAGAGCGCATGGA 960
Db 901 CAGATTTTATGAGGCTCTGAGAACTCAAGATTAAGTTAGTGAAGAGCGCATGGA 960
Qy 961 AGTATCCGAAATCTTCTTATGATGATGAGGATATATTTCCATGTTTCTGCTGTA 1020
Db 961 AGTATCCGAAATCTTCTTATGATGATGAGGATATATTTCCATGTTTCTGCTGTA 1020
Qy 1021 TGGTGAATCTTAAAGCCTCTCTCTGAGATATTTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 TGGTGAATCTTAAAGCCTCTCTCTGAGATATTTGCTGCTGCTGCTGCTGCTGCTG 1080
Qy 1081 GTGATTCACCAAGATTAAGCTGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1081 GTGATTCACCAAGATTAAGCTGATGATGATGATGATGATGATGATGATGATGAT 1140

Db 1081 GTGATTCACCAAGATTAAGCTGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 1141 ATTATCAGATGATTAATATCTTTCAGTAAATTAATGAAACAAATAGCTTGACAG 1200
Db 1141 ATTATCAGATGATTAATATCTTTCAGTAAATTAATGAAACAAATAGCTTGACAG 1200
Qy 1201 CTCACCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db 1201 CTCACCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Qy 1261 ACTGATTAATAA 1271
Db 1261 ACTGATTAATAA 1271
RESULT 2
AX093580 1271 bp DNA linear PAT 30-MAR-2001
LOCUS
DEFINITION
Sequence 1 from Patent WO0118190.
ACCESSION
AX093580
VERSION
AX093580.1 GI:13510017
KEYWORDS
SOURCE
Solanum tuberosum (potato)
ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
1 Ehrhardt, T., Lerchl, J., Stitt Nigel, M., Zrenner, R. and Schroeder, M.
Dihydrochalcone extracted from plants
Patent: WO 0118190-A 1 15-MAR-2001;
BASF AKTIENGESELLSCHAFT (DE)
FEATURES
source
1..1271
/organism="Solanum tuberosum"
/mol_type="unassigned DNA"
/db_xref="taxon:4113"
9..1049
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC35419.1"
/db_xref="GI:13510018"
/translation="MELSTIHPDDPHLHRRDDVLLKAVVSHSHHFGRAIIVNPMLKPP
ITTTAAVAVYRAILKSLPVSDSPNPMLTLYLDTSPWEIKLARSQVFFVKLYPA
GATPNSODGVTDLFGKCLPYLQBMWHNNPLVHBEVNPVDMREKVPFJETVLRP
LVKRFQQLKAVHEHTTIDAVKRVESCTEGFVAATVTPQHLVAINRNSLFQGLJOPHY
CLFVLREAIHREALVSAVTSGRKRFELGDSAPHRRRREKSCGAGIYNAPVALSVY
AKVFEKNALDKLEAFSTFNPDPFGLPRNNSKIKLSKTPMKVPESFVYASGDIIPMF
AGEMLDWLBPAL"
ORIGIN
Query Match 100.0%; Score 1271; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 4.8e-313;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGCAAAATGAGAGCTCTCAATACACACCTGAGATGGCATCTTCACTCCCGTAGTG 60
Db 1 TTGCAAAATGAGAGCTCTCAATACACACCTGAGATGGCATCTTCACTCCCGTAGTG 60
Qy 61 GTATGTTCTTAAAGCAGTTGTCTCTCAGAGTGCAATCATCTTGGAGGGCAATATGCA 120
Db 61 GTATGTTCTTAAAGCAGTTGTCTCTCAGAGTGCAATCATCTTGGAGGGCAATATGCA 120
Qy 121 TGGCAAAATTTGAAGCCTCTATCACTACACCTGCTGCTGTAGCAATACCGGAGGCGA 180
Db 121 TGGCAAAATTTGAAGCCTCTATCACTACACCTGCTGCTGTAGCAATACCGGAGGCGA 180
Qy 181 TATGAAATCTTAACTGCTGTAGTAGATTTCAACCTCTTATAGCACTTATTTAGAG 240
Db 181 TATGAAATCTTAACTGCTGTAGTAGATTTCAACCTCTTATAGCACTTATTTAGAG 240
Qy 241 ATACAAACGAGCTCTATGGAATCAATACGAAAGAGAGCGAGTGTATTTGGGGTGA 300
Db 241 ATACAAACGAGCTCTATGGAATCAATACGAAAGAGAGCGAGTGTATTTGGGGTGA 300

Db 241 ATACAAACAGTCTCTATGAGAAATCAAACTAGCAAGAGAGACCAAGTCGTATTTGGGGTGA 300
QY 301 AGTTGTAACCTGCTGCTGTCGACACACAATTTCTCAAGATGAGTACTGATCTTTTGGGA 360
Db 301 AGTTGTAACCTGCTGCTGTCGACACACAATTTCTCAAGATGAGTACTGATCTTTTGGGA 360
QY 361 AGTGTTCACAGTTTCTCAAGAAATGTTGACATTAATATGCTCTGCTGCTTCAATGAG 420
Db 361 AGTGTTCACAGTTTCTCAAGAAATGTTGACATTAATATGCTCTGCTGCTTCAATGAG 420
QY 421 AGGTTCATATCCTGAGGTTGACATGTTGATGAGAAAAGTATTCATGAAACGGTTC 480
Db 421 AGGTTCATATCCTGAGGTTGACATGTTGATGAGAAAAGTATTCATGAAACGGTTC 480
QY 481 TAAAGCCGTGTGTCAGAAATTTCCAAATGGAAGTCTGATGAGACATGTTAACCA 540
Db 481 TAAAGCCGTGTGTCAGAAATTTCCAAATGGAAGTCTGATGAGACATGTTAACCA 540
QY 541 TTGATGCTGTTAAGTTTGTGGAATCTTGACCTGAAAGATTTTGTGACAGCAACTGTGACC 600
Db 541 TTGATGCTGTTAAGTTTGTGGAATCTTGACCTGAAAGATTTTGTGACAGCAACTGTGACC 600
QY 601 CACAACATCTGTTTGAACAGAAATCTCTCTTCCAAAGGGGGCTTACAAACCGCATAT 660
Db 601 CACAACATCTGTTTGAACAGAAATCTCTCTTCCAAAGGGGGCTTACAAACCGCATAT 660
QY 661 ACTGCTTCCAGTCTCTCAAAAGAGATCCACAGGAGGCACTTGTCTAGCTGTAA 720
Db 661 ACTGCTTCCAGTCTCTCAAAAGAGATCCACAGGAGGCACTTGTCTAGCTGTAA 720
QY 721 GTGGAAGTAAAGATTTTCTTGGGACTGATAGTCTCTCATGATAGACGAAAGAA 780
Db 721 GTGGAAGTAAAGATTTTCTTGGGACTGATAGTCTCTCATGATAGACGAAAGAA 780
QY 781 AGGTTCTTGTGATGATGCTGTATTTAACAATGACCTGAGCTTGTACATATATG 840
Db 781 AGGTTCTTGTGATGATGCTGTATTTAACAATGACCTGAGCTTGTACATATATG 840
QY 841 AGGTTTGAAGAAAGAAATGCACTGCAAGCTTGAAGCACTTCACTAGCTTCAATG 900
Db 841 AGGTTTGAAGAAAGAAATGCACTGCAAGCTTGAAGCACTTCACTAGCTTCAATG 900
QY 901 CAGATTTTATGAGCTTCTTCAAGAACTCAAGATTAAGTTAGTAAAGCCGATG 960
Db 901 CAGATTTTATGAGCTTCTTCAAGAACTCAAGATTAAGTTAGTAAAGCCGATG 960
QY 961 AGGTAACCGAATCTTTTCTTATGATCAGAGATTAATTCATGTTGCTGTG 1020
Db 961 AGGTAACCGAATCTTTTCTTATGATCAGAGATTAATTCATGTTGCTGTG 1020
QY 1021 TGCTGACTGCTGCTGCTCTCTCTGAGATCATTTGATCTTGTACTGTATAT 1080
Db 1021 TGCTGACTGCTGCTGCTCTCTCTGAGATCATTTGATCTTGTACTGTATAT 1080
QY 1081 GTGATTCACCAAGATATAGACTGTAGGTATCATTTTCTTCACTGTTGAT 1140
Db 1081 GTGATTCACCAAGATATAGACTGTAGGTATCATTTTCTTCACTGTTGAT 1140
QY 1141 ATTATCAGATGATATATCTTTTCACTAATTAATTAAGAAACAATAGCTTTCACG 1200
Db 1141 ATTATCAGATGATATATCTTTTCACTAATTAATTAAGAAACAATAGCTTTCACG 1200
QY 1201 CTCACCAAGTCTCTGATTTCTGAAAGTCTTAAATTTGTTGATTTTGAAGAT 1260
Db 1201 CTCACCAAGTCTCTGATTTCTGAAAGTCTTAAATTTGTTGATTTTGAAGAT 1260
QY 1261 ACTGATATAAA 1271
Db 1261 ACTGATATAAA 1271

LOCUS AX470409 1743 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 2 from Patent WO2052012.
ACCESSION AX470409
VERSION AX470409.1 GI:22205539
KEYWORDS
SOURCE
ORGANISM
Beta vulgaris
Beta vulgaris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Beta.
REFERENCE
1 Kanhonou, R.A., serrano Salom, R. and ros Palau, R.
AUTHORS
Sugar beet genes involved in stress tolerance
JOURNAL
Patent: WO 02052012-A 2 04-JUL-2002;
Cropdesign N.V. (BE)
FEATURES
Location/Qualifiers
1..1743
/organism="Beta vulgaris"
/mol_type="unassigned DNA"
/db_xref="taxon:161934"
ORIGIN
Query Match 50.8%; Score 646; DB 6; Length 1743;
Best Local Similarity 76.4%; Pred. No. 1e-153;
Matches 793; Conservative 0; Mismatches 245; Indels 0; Gaps 0;
QY 3 GCAAAATGAGCTCTCAATCAACAACCTGATGATTTGGATCTTCCGATG 62
Db 193 GTAAAGATGAGACTCTTCAACAGCCCTGATGACCTGATCTTCACTCCCGATGA 252
QY 63 GATGTTCTTAAAGCAGTGTCTCTCAAGGACATCACTTGGAGGCAATAGTCATG 122
Db 253 GATCTTCTGCTGCTGCTGCTCTCTCAAGGACATCACTTGGAGGCAATAGTCATG 312
QY 123 CCAATTTGAAGCTCTCTCACTACCACTGCTGCTGTAGCATACCGGAGCGATA 182
Db 313 CCGATTAAGGCTCTCTGTTACTACACAGGAGGCTGATTTGCCAAAGTCTAT 372
QY 183 TTGAATCTTAACTGTTGATGATTTCAACCTCTTATGACATTTATTTGACAT 242
Db 373 ATGAAGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
QY 243 ACAACGATCTTAAAGAAATCAAACTAGCAAGAGAGACGATGATTTGGGGTGA 302
Db 433 ACAGCCAGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
QY 303 TTGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
Db 493 TTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
QY 363 TGTTTACAGTTTCTCAAGAAATGTTGATGATTTGCTCTGCTGCTGCTGCTGCTG 422
Db 553 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 612
QY 423 GTTATCTATCTGAGGTTGATGATTTGATGATGATGATGATGATGATGATGATGAT 482
Db 613 GTTACAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 672
QY 483 AGACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
Db 673 AGACCTTAAATTCAGAAATTAACACAGCTTAAAGTTGATGATGATGATGATGAT 732
QY 543 GATGCTGTTAAGTTTGTGATTTGATGATTTGATGATTTGATGATTTGATGATTT 602
Db 733 GATGCTGTTAAGTTTGTGATTTGATGATTTGATGATTTGATGATTTGATGATTT 792
QY 603 CAATCTGTTTGAACAGGAATCTCTTCAAGGGGGCTTACAAACCGCATATATAC 662
Db 793 CAGCACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
QY 663 TGCTTCAGTCTCAAAAGAGATCCACAGGAGGCACTTGTGATGATGATGATGAT 722
Db 853 TGCTTCAGTCTCAAAAGAGAAATCAATGACAGGCACTTGTGATGATGATGATGAT 912

QY	723	GGAAAGTAAAGATTTTCTTGGAGCTGATATGCTCTCTCATGATATGACGAAGAAAGAG	782
Db	913	GGAGACCAAGCATTATTTCTTGAGACTGATAGTCTCCATCAAGAAAGGGGAGAAAGAA	972
QY	783	TGTTCTTGATGATGCTGCTGATATTTACAATGCACTGTAGCCTTGCTGATATATGCAAG	842
Db	973	TGTTGCTGAGATGCTGGAAATCTATAATTTCCCTGTTGCTCTATCACTATATGCAAA	1032
QY	843	GTGTTGAAAGAAAGAAATGCACTGCAACAAGCTGAAGACTTACATAGCTTGAATGAGCA	902
Db	1033	GTAATTGAAGAGGCTGCTGCTGATCCCTTGACAGTTAGAGCATTTTACAAAGCTTTAATGACCC	1092
QY	903	GATTTTATGGCTTCTTAGAACAACCTCAAAATTTAATGATAGTAAGACGCATGGAAG	962
Db	1093	GATTTCTATGCTTCCAAAGAAATTCGTGAAGATCAAGATGAAAGAAAGAAACCATGGA	1152
QY	963	GTACCCGATCCTTTCTTATGCAATCAGAGATATTTATCCAGTTGCTGCTGATAG	1022
Db	1153	GTCCTAGAGCGATATCCTTTCCCATCCGAGAAATTAATCCCTATGTTTGTCTGACAAATG	1212
QY	1023	CTGACATGCTGCCGCT	1040
Db	1213	CTTGACTGAAGACATCT	1230

QY 1029 TGCTGCCGCT 1040
DB 1021 TGAAGCATCT 1032

RESULT 5
AF0000146
LOCUS Arabidopsis thaliana dihydroorotase (pyr4) mRNA, complete cds.
DEFINITION AF0000146
VERSION AF0000146.1 GI:2121272
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1362)
Zhou, L., Lacroix, F. and Thornburg, R.
Characterization of the Arabidopsis thaliana cDNA encoding
Dihydroorotase (Accession No. AF0000146) (PGR97-115)
Plant Physiol. 114, 1569 (1997)
2 (bases 1 to 1362)
Zhou, L., Lacroix, F. and Thornburg, R.W.
Direct Submission
Submitted (16-APR-1997) Biochemistry and Biophysics, Iowa State
University, Ames, IA 50011, USA
Location/Qualifiers
1..1362
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
<1..>1362
/gene="pyr4"
86..1219
/gene="pyr4"
/EC_number="3.5.2.3"
/function="catalyzes the intramolecular condensation of
carbamoyl aspartate to form dihydroorotase"
/codon_start=1
/product="dihydroorotase"
/protein_id="AA071134.1"
/db_xref="GI:2121273"
/translation="MIRKLVSPYSGFSQKIKFRSSEKVTAVRMELITPDPMH
LHRDGLHVVHSHSNFKRAIVMPKLPVSTAAIYKFKIKALPSESSFDP
LMTLYLDKTLPEIRLARRESGVYAVLYPAGATNSODGVDLFGCLPVEEMK
QNMPLVHGEVTDPSIDVREDEKIETIVLQPLORLPOLKVMHETMDVAVNEVS
CKESVQATVTPQHLNRNLFGGLOPHNYCPULKREIHRRAIYKAVTSKKEFP
LGTDSAPHERKSSCGCAGIYSAPIALSLYAKVFDGALDLKLEAFISNPFDFYG
LPRNSKITLKSSPMKVDPVDFPFGELVPMFAGETLQWOLK"
polyA_signal
1325..1330
/gene="pyr4"

ORIGIN
Query Match 46.2%; Score 587; DB 8; Length 1362;
Best Local Similarity 72.9%; Pred. No. 1.2e-138;
Matches 755; Conservative 0; Mismatches 280; Indels 0; Gaps 0;

QY 3 GCAAAATGAGCTTCATACACAACTGATGATGCTTCATCTCCGATGATG 62
DB 176 GTAAGAAAGAGCTACACATCACTAACCTGATGATGATGCTTCATCTCCGATGATG 235
QY 63 GATGTTCTTAAGGAGCTGCTCTCAAGTGCATCACTTTGGAGGCAATAGTCATG 122
DB 236 GATCTTCTTCAATGCTGTTGTTCCCAAGTGCATGATTTTAAAGAGCATTTGTATG 295
QY 123 CCAATTTGAAGCTCTCTATCACTACCACTGCTGCTGTGATGCAATACCGGAGCGCAT 182
DB 296 CCAATTTGAAGCTCTCTATCACTACCACTGCTGCTGTGATGCAATACCGGAGCGCAT 355
QY 133 TTGAATCTTACCTGTTGATGATGATTTCAACCGCTTATGACATTTATTTGACAGAT 242
DB 356 ATGAAGCTTTCCTGATGAGATGATGATTTGATTCATTTATTTATTTGATGAC 415

QY 243 ACAACAGTCTATGGAATCAAACTAGCAGAGAGCCAGTCTATTTGGGTGAAG 302
DB 416 AAACTCTACCTGAGAGATCAGGCTTGCAGAGGAAAGTGTGTTATGCGGTGAAG 475
QY 303 TTGACCTCTGCTGTGTCACAGCAAAATCTCAAGATGAGTGAATCTTTTGGGAG 362
DB 476 CTGTACCTCTGCGGAGCAACCACTCTCAAGATGAGTGAATCTTTTGGGAG 535
QY 363 TGTTTACAGTTTACAGAAATGTTGAGCATATATGCTCTGCTGTTTATGAGAG 422
DB 536 TGCTTACCGGTACTAGAGAGATGTCACAAAGATGCTTGTGCTGATGAGGAG 595
QY 423 GTTACTAATCTGAGGTGATCATGTTGTATGAGAAAGATATTCAATGAAAGGTTCTA 482
DB 596 GTACAGATTCAGATTCGATGCTTTTTCACCGGAGAAATCTTCAATGAGAGTTCTG 655
QY 483 AGACCGTGTGAGCAAAATTTCCAAATGAGGTGATGAGAGATGTTACCACTT 542
DB 656 CAGCTCTAATCCAGCGCTTCCGAGCTGAAGTGTATGAGACATCACACACATG 715
QY 543 GATGCTGTTAAGTTGTTGATTTGATCTTGACATGAAGATTTGTCAGCACTGACCCCA 602
DB 716 GATGCTGTTAAGTTGTTGATTTGATCTTGCAAAAGAGGTGCTGATGAGATGACACCA 775
QY 603 CAACATCTTCTTTCACAGAGAAATCTCTCTTCCAAAGGGGCTTCAACCGCATTAATAC 662
DB 776 CAACATCTTCTTTCACAGAGAAATCTCTCTTCCAAAGGGGCTTCAACCGCATTAATAC 835
QY 663 TGCTTTCAGTCTCTCAAAAGAGATCCACAGGAGGCACTTGTCTGATGATGATG 722
DB 836 TGCTTTCAGTCTCTCAAAAGAGATCCACAGGAGGCACTTGTCTGATGATGATG 895
QY 723 GGAAGTAAAGATTTTCTTGGAGCTGATAGTCTCTCATGATTAACAGAAAGAG 782
DB 896 GGAAGTAAAGATTTTCTTGGAGCTGATAGTCTCTCATGATTAACAGAAAGAG 955
QY 783 TGCTTTCAGTCTCTCAAAAGAGATCCACAGGAGGCACTTGTCTGATGATGATG 842
DB 956 TGCTTTCAGTCTCTCAAAAGAGATCCACAGGAGGCACTTGTCTGATGATGATG 915
QY 843 GTGTTGAAAGAGAAATGCACTGACAGAGGCACTTGTCTGATGATGATGATG 902
DB 1016 GTCTTTATAGAGGCGGCTTCCGAGCTGACAGATGGAACCTTTCACAGTTCAATGAGACT 1075
QY 903 GATTTTATGAGGCTTCTTGAAGACACTCAAGATTAAGTTAGTTAAGAGCCATGAG 962
DB 1076 GATTTTATGAGGCTTCTTGAAGACACTCAAGATTAAGTTAAGTTAAGAGCCATGAG 1135
QY 963 GTACCCGATCTTTCTTATGATCAAGAGATTTATCCAGTTTGTGCTGAGAAATG 1022
DB 1136 GTTCCAGAGCTTTTCAACCTTCCATGAGAGATGCTCTTATGTTGCTGAGAAATG 1195
QY 1023 CTGAGCTGTTGCCG 1037
DB 1196 CTTCAATGACACCG 1210

RESULT 6
AK122010
LOCUS AK122010
DEFINITION AK122010 1750 bp mRNA linear PLN 29-OCT-2003
Oryza sativa (japonica cultivar-group) cDNA clone:J033111101, full
insert sequence.
AK122010
VERSION AK122010.1 GI:37991633
KEYWORDS FLI cDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharatoidae; Oryzaceae; Oryza.
REFERENCE
1 The Rice Full-Length cDNA Consortium, National Institute of

Db	1105	ACAAATCTTCTTGTAATGACAGTGTCTCTCTATGATTAACGAGAAAGAAATGCTCTTG	1164	CDS	/gene="slr0294" 1..480 /gene="slr0294" /note="ORF_id:slr0294 unknown protein" /codon_start=1 /transl_table=11 /protein_id="BA18517.1" /db_xref="GI:1653605" /translation="MTKRALIFPAQGLGWTAKDAERAYQALNVNPSDLOSLDYKGE YVIEDVTDIDIVLSVLADPAGVLLDRQKQAOQGVTKKAKIKETBQSHQEI ADFKSFHQSHSLMELLIKVYSTAKRFGLKEPWIEMIDAYEAKIGTDSQVA" 428..823 /gene="slr0392" 428..823 /gene="slr0392" /note="ORF_id:slr0392 unknown protein" /codon_start=1 /transl_table=11 /protein_id="BA18518.1" /db_xref="GI:1653606" /translation="MLMRQKCLITRKQLELGLTSYQAQRTKELKPISKTKGKN YNIQEVILSINTLENKRIQKRTSLKQVLDILPLDNIIPISFTQENTSELQDLA QALTTTLEVNRLAKLERRSLKIGRYE" 816..1181 /gene="slr0393" 816..1181 /gene="slr0393" /note="ORF_id:slr0393 unknown protein" /codon_start=1 /transl_table=11 /protein_id="BA18519.1" /db_xref="GI:1653607" /translation="MNEDLDNDFOQKYEKYKLAARRKCHDLKLRREENIRGYTPR NVFNRRDSTIEGKKMONOYKQKACFCGAFVQSPYLEIDHIKPIATHPBELSV NITKULDLCPCKRKRGKGR" complement(1465..1884) /gene="sl10364" complement(1465..1884) /gene="sl10364" /note="ORF_id:sl10364 unknown protein" /codon_start=1 /transl_table=11 /protein_id="BA18520.1" /db_xref="GI:1653608" /translation="MLPPTMESPMKTSNDRILITLQODYANFPWHTFYSIDPVNYFQ DPLSQFRLKRYEMIDFLGRMQRADIDQHEITQVNOCELITRTLTMTSPLPKPRI AIGCSSELTINLEGLISHIDYDCSPWDVYKQHSR" 2013..3137 /gene="pgk" 2013..3137 /gene="pgk" /note="ORF_id:slr0394" /codon_start=1 /transl_table=11 /product="phosphoglycerate kinase" /protein_id="BA18521.1" /db_xref="GI:1653609" /translation="MPPDNGSITDDPTIRPALPTIKDLSKGAIVILGSHRGRPKGV VDSMRITPVGRLDELIGQPVKCDDCIGAEVTAKIKSLPRNGVALLENIPFAGBERG NDAPFAALANALLVNDAPGTAHRAHASTEGVTHLSPVAVGLTEKELQFLOGAL EAPRPLVAIVGSKSVSKIGVITLIDKCDKLIIGGGMITFYKAGQWNLGSLVES DKLDLAKLMAKKEKVEFLPTDVVADAFADANAQTVGAIDPGWNLGSLVGP SVKTFODALAGCGTVINMGVGFEPFPAVGTGAICSLAELTASGTVIIIGGSDSV AAVKKVGAEXMSHISTGCGASLELLSGKVLPGIALDDR" 3406..4335 /gene="ntcB" 3406..4335 /gene="ntcB" /note="ORF_id:slr0395"
Qy	791	TGGATGTGCTGATTTACATGACCTGTAGCTTGTGAGTATATGCGAAGGTGTTGA	850	gene	/gene="slr0395"
Db	1165	TGGTTGTGAGAAATATATAGCGCTCTGTGGCTCTTTCCTTATGCGAAAGTATTCGA	1224	CDS	/gene="slr0395"
Qy	851	AAAGAAATGCACTGACCAAGCTTGAAGCATTCAGTCTCAATGACCAATTTTAA	910	gene	/gene="slr0395"
Db	1225	ACAGGCCGGGTCTTGATTAAGCTAGAACGATTTACAAAGTCCCAATGGCCAGATTTTAA	1284	gene	/gene="slr0395"
Qy	911	TGGGCTTCTTGAAGAACCACTCAAGATTTAGTGAAGAACGCAATGGAAGTACCCGA	970	CDS	/gene="slr0395"
Db	1285	TGGCTTTCGAAGAACCTTCAAGATGTCTTGAAGAACGCGCTGGAAGTACTGA	1344	gene	/gene="slr0395"
Qy	971	ATCCTTTCTTATGATCATGAGAGATTTATTTCCATGTTTGTGTGTAAGTCTGACTG	1030	CDS	/gene="slr0395"
Db	1345	TACATACTCATATAGTTACAGAGAGATCGGCTATGTTTACAGGCAACACCTCGAATG	1404	gene	/gene="slr0395"
Qy	1031	GTTCCGGGCTCTCTGTGAGAA	1052	gene	/gene="slr0395"
Db	1405	GCTTCATCCGATCACTGAA	1426	gene	/gene="slr0395"
RESULT 7					
D90915		130001 bp	DNA	linear	BCT 05-JUN-2001
LOCUS					
DEFINITION					Synechocystis sp. PCC 6803 DNA, complete genome, section:17/27,
ACCESSION					D90915 AB001339 BA000022
VERSION					D90915.1 GI:1653604
KEYWORDS					Synechocystis sp. PCC 6803
SOURCE					Synechocystis sp. PCC 6803
ORGANISM					Bacteria; Cyanobacteria; Chromococcales; Synechocystis.
REFERENCE					
AUTHORS					1 Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Sazuka, T., Miyajima, N., Sugiyama, M. and Tabata, S. 2 Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosewa, M., Sugiyama, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Nanno, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S. Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64k to 92k of the genome DNA Res. 2 (4), 153-166 (1995) 96127529 8590279
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS
JOURNAL					gene
MEDLINE					CDS
PUBMED					gene
REFERENCE					CDS
AUTHORS					gene
JOURNAL					CDS
MEDLINE					gene
PUBMED					CDS
REFERENCE					gene
AUTHORS					CDS

gene
CDS
/codon_start=1
/transl_table=11
/product="transcriptional activator protein; Ntcb"
/protein_id="BAA18522.1"
/db_xref="GI:1653610"
/translation="MRLEQLOFLKVAELSGFOOALQESHTYSTISROITQGLSEALM
COLFHGAQAKLTAVAGEFLRPARKICQEDVDASELSLPGQKQSLDVLVAIHVC
SLSPSLPFCGHPQIOLRVTALGSRDLAKVLDGVLVDALIMSHRINTKKEALH
PLVEQICGLIMASHDPLTKKFTFMBNLGFPVITFDGVRKMRILVEFSSREPLPL
VSELENIPEAPGVGVGSBMLALMPSQSLVTPVDNPNFSVRLFCPESGDQDFRRQ
SVYTVDRQLQIPPVAFEFNVLVDHRCGALTVC"
complement(4355. .6988)
/gene="alac"
complement(4355. .6988)
/gene="alac"
/note="ORF ID:s110362
gene:alac or 1ovb"
/codon_start=1
/transl_table=11
/product="alanyl-tRNA synthetase"
/protein_id="BAA18523.1"
/db_xref="GI:1653611"
/translation="MTTTPVLSGPRIEQRPOPLNFPAQDQHOLIPASGLVEPDPVLT
IAGMPEPFIQLQGSAPPRATTSSQKIRTDINVRGTAAHRTFFEMIGNSFGDY
FSQSLAAWMLSTQVFKLPAERLIVSYVEEDDEAFIMRDEIGIPARIQKNGADNP
FVWSPGTPQPSCEIYYDPHPRLGDKELDSDSRFLIEFNVLVPMQYNDNNGNLP
LEKRNIDQKGLERMAQITLQKQNNYETDLIPPIQTNANTAGIDVQAQNEKTVSLK
VICHYVSVVHMLADGISNLGRGYVLRLLIRVRHGRLLAGINGEFTTKVAATVO
LAQFPIPVABRSQSLLEQELQREBEAFLKTLTERGSKLADIMADGVTEIAGDAPLTY
DTGPEPLTQELIAEBOGILVDEGEKAMQOQRSKAHEITDLTVOSLKLXNH
IHPEPFIAGYDLOSALIVKVAIVGSELVQVAVAGTQVIVIDQPPYSGEGQIGKG
FLNDEPDLILIEDYKRESGFIHFGKRGVTCQITTTATITDACKRRQAQNAHTH
LALQALRVDEGDISQSGSLVDPNLRIPDPSPRATYMEELQQLIEDLINWIAEAQTH
EVAMPPLADKAKGALAMFGKCYCAVYIVDPGVSJELCGTHVANTAEIGFKTVA
QTKQLEAKGRIEAVAGPSVLDIYNREAVNVEIDGRUKAPEELIPRWHLQQLKAS
AVNTASIPESGKSVLVAASLPQLVTKTQFKAIVLDTIVGESLKTAAERLQKCES
DASLTLPALTAQTLTAAELG"
complement(7169. .7639)
/gene="s110361"
complement(7169. .7639)
/gene="s110361"
/note="ORF ID:s110361
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAA18524.1"
/db_xref="GI:1653612"
/translation="MPMSPSVFTGEPFLRRMLGMSGVITMLANQGLSCGASKAQD
CRQLILITQMAEKGAKRYRSTTEVILNMAADAPATKTRTEINLEDDPVAQNGEL
VSYIQGNAATRTYKALSKEDILTLAQLAQKQVATTGGQEDQVITINVHCQP"
complement(7721. .8554)
/gene="s110360"
complement(7721. .8554)
/gene="s110360"
/note="ORF ID:s110360
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAA18525.1"
/db_xref="GI:1653613"
/translation="MENIMKEFWQPMWGLKLIIVFLPSMLVLMYPIVALVSGKRLAMQW
QAVTVRQQLVAVSLVYSPSLPIWATWMLDQSLDYGSLVSDQLSLIAIVANVF
GIGLTLAQGGIGILQWHRKEMWDLVLSPLFLITALLAIWVBEILIRGIFQNDQD
LFWQCAATVSTIPIAILHLMERRLTYAQLPGMILGNVILMAELDNGNIALAMGIG
GGMVFLAAVATGAILITVPTNDGNIWVGKANGPLAGLAGITICLLTLLATLALPPLHSP
L"
complement(8742. .9209)
/gene="s110359"
complement(8742. .9209)
/gene="s110359"
/note="ORF ID:s110359
unknown protein"

Query Match	Best Local Similarity	27.2%; Score 345.6; DB 1; Length 130001;
Matches 608; Conservative	0; Mismatches 389; Indels 9; Gaps 2	
Query	13	AGCTCATACACACAACTGATATTTGGCATCTTCACTCCGTGATGATGATCTTCA 72
Db	20191	AGCTTACCATCATCTCGACCCGACACACGCGACCTGCATCTACCGGATGAGCGCCCTGA 202
Query	73	AGCGAGTTGTCTTCACAGTGCACATCTCTTGGAGGAGGAGCAATGATCAGCCAAATTTGA 132
Db	20251	AAGCAGTTCTACCCCACTGTGCGTCAATTTCTAGCGGATGATGATCAGCCAAATTTGA 203
Query	133	AGCCCCCATACACACACACTGCTGCTGCTGTAGCATACCGGAGGCGATATTTGAATCTT 192
Db	20311	AGCTCCGCTGCGTTCCTGCGGTGAGTGGCCGCCCTATGCGGAAAGGATTTCCGCCGCA 203
Query	193	TACCTGTGATAGTATTTTCAACCTCTTATGACATTTATTTTGAACAGTAAACACAGTC 252
Db	20371	TACCTGCGGAGGCGCAGTTTGAAACCGCTGATACCTTTATCTAACGGATTAATCTAATC 2043
Query	253	CTATGGAATATCAAACTAGCAAGAGAGAGCCAGCTGTATTTGGGTGAAATTTGACCCCTG 312
Db	20431	CCGAGAAATTAATTTGGCGCCAAAGCAATCCCAATTTGTTAAAGCATTAATAATCTATCCG 20454
Query	313	CTGGTGCACAGCAAAATTTCTCAAGATGAGTACGATCTTTGGGAGAGTGTTAACAG 372
Db	20491	CTGAGGACACCAATTTCTGACTTTGGGTGACGSA---TATCATGATGTGATGCG 2054
Query	373	TTCTACAAGAAATGCTGTGAGCATTAATATGCTCTGCTGCTGATCAGAGAGTTTACTATC 432
Db	20548	TGTTGGCAGCATGAGAAACAGTGTGACTTGCCTTATTACTCAGCGGACAGTAAACGACA 2060
Query	433	CTGAGTTGACATGTTTATAGAGAAAGATTAATCATTTGAAACGGTTCTAAGACGGTTGG 492
Db	20608	GTAATATTTGATATTTTTCACCGAGAAAGATTAATCATTTGAAATTAATTTATTTCTCTCA 2066
Query	493	TGCAGAAATTTTCACAATTTGAAGGTCGTGATGAGAGCATGTACCAACATGATGCTGTTA 552
Db	20668	GGGAAAAATTTCCCGGCTGGCGGTGATGTGCGAACAATTAATCTACATTCAGATGCGATGC 20722
Query	553	AGTTTGTGAATCTTGCACTGAGAGATTTGTTGAGCACTGTACACCCCAACATCTTG 612
Db	20728	AATTTGTTTATCTGTCAATAA-----TATTCAGCGCATATTAAGCCCAACATTTAC 2078
Query	613	TTTTGAACGAAATTTCTCTTTCACAGGGGCTTCAACCGGCATTAATTTACTGCTTCAG 672
Db	20782	TATTTAGTCGACAGCTTTATTTCAAAAGTGTGCAATTTGTCCCACTTTTATGCTTCGCAA 20844
Query	673	TCCCTAAAGAGAAATCAACAGGAGGCACTTGTGTACGTGTACAGTGAAGTAAATA 732
Db	20842	TTCTGAACCGGAGAAACATCTGTTTATCTATGCTTACACCGGCGCATCTGGTAATCTTA 20901
Query	733	GATTTTCTTGGACATAGTGTCTCTCATATATGAGAGAAAGAAAGTGTCTTG 792
Db	20902	AATTTTCTTCGGACATGATGTGTCTCCCAAGCTGCAATGTAAAGAAATTTATGTG 20961
Query	793	GATGTGCTGTATTTACATATGACCTGTAGCTTGTCAATATATGCAAGTGTTTGAAA 852
Db	20962	GCTGTGCTGTGTCTATTTCTGCTTCATATGATGAGCTATATATCCCAAGCTTTTGGAAA 21021
Query	853	AGGAAAATGCACTCGAACCTTGAAGCATTTACATAGCTTCAATGACAAATTTTATG 912
Db	21022	GTTGTATATTCCTGAGATTAAGTAAAGCTTTGCTAGCTTTTATAGCCAAAGATTTTAAAC 21081
Query	913	GGCTCTTAGAACCACTAAAGATTAAGTTAGTAAACGCAATGAAAGTATCCCGAAT 972

Db 21082 AATTGCCCGCTATACGACGAAATTACGTTGATGAAACCCCTGGCGATTCTCGTGG 21141

Qy 973 CCTTTCTTATGCATCAGAGATATTAATCCATGTTTGTGTGA 1018

Db 21142 AATTACCTTTCCCTGAATCGATGTGGTGCCTCGCGGCGGCGGA 21187

RESULT 8
BX321858 306050 bp DNA linear BCT 23-APR-2003
LOCUS Nitrosomonas europaea ATCC 19718, complete genome; segment 3/10.
ACCSSION BX321858 AL954747
VERSION BX321858.1 GI:30180174
KEYWORDS complete genome.
SOURCE Nitrosomonas europaea ATCC 19718
ORGANISM Nitrosomonas europaea ATCC 19718
Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
Nitrosomonadaceae; Nitrosomonas.

REFERENCE
AUTHORS 1 Chain, P., Lamerdin, J., Larimer, F., Regala, W., Land, M., Hauser, L.,
Hooper, A., Klotz, M., Norton, J., Sayavedra-Soto, L., Arciero, D.,
Hommers, N., Whitaker, M. and Arp, D.
TITLE Complete Genome Sequence of the Ammonia-Oxidizing Bacterium and
JOURNAL Obligate Chemolithoautotroph Nitrosomonas europaea
PUBMED 12700255
REFERENCE 2 (bases 1 to 306050)
AUTHORS Larimer, F.
TITLE Direct Submission
JOURNAL Submitted (12-NOV-2002) Submitted on behalf of the Nitrosomonas
genome consortium, the DOE Joint Genome Institute, Production
Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598,
USA, and the Genome Analysis Group, Oak Ridge National Laboratory,
1060 Commerce Park Drive, Oak Ridge, TN 37831, USA;
larimerf@ornl.gov
Nitrosomonas genome consortium
Location/Qualifiers
1. 306050
/organism="Nitrosomonas europaea ATCC 19718"
/mol_type="Genomic DNA"
/strain="ATCC 19718"
/db_xref="taxon:228410"
224. 1471
/locus_tag="NE0566"
224. 1471
/locus_tag="NE0566"
/function="InterPro IPR001633:IPR001789 COG5 COG2200"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="Domain of unknown function 2:Response regulator
receptor domain"
/protein_id="CAD84477.1"
/db_xref="GI:30180175"
/translation="WKADIFTRILLDDPFILKLMKMLAMGFSSVTCENGVA
VKLLDVAATRPDLILLDMKMGDIEVRALVERKSGSLIVSGEDERLRTTEKL
VQAHRIPIGLYLOKVPQESLRMLKLDVSSWRSGERNRRIKYADATTITADISQO
LINYQPKVRVDVTVAGVAVLVAMHNLVDQVYEDPFIVGAENGLGRITHTVLSR
AFIAQKWQSGYELKVAVNVSMNLTLSDODAVVRLAAEAILPGKIVITEESOL
MGDRAVSEIITRLRIKRFPHISIDDFGCHSLQRLRIPEDELKIDGGFPHRASGDE
TLRAIYDASLALAKOLNMEAVAEQVEDREDELLRQTCDLAQGNFISRPALALGS
WMEKWHVRIEASLIVSPKAKAL"
complement (1514. .2062)
/gene="dfrA, tmp"
/locus_tag="NE0567"
complement (1514. .2062)
/gene="dfrA, tmp"
/locus_tag="NE0567"
/complement (1514. .2062)
/gene="dfrA, tmp"
/locus_tag="NE0567"
/BC_number="1.5.1.3"
/function="InterPro IPR001796 COG5 COG0262"
/codon_start=1
/evidence=not_experimental
/transl_table=11

gene
CDS
/product="Dihydrofolate reductase"
/protein_id="CAD84478.1"
/db_xref="GI:30180176"
/translation="MTIPESLSPRLAIIAAVSAANVIGIINTLPMHLPADLKHFKQLT
TQOIVMGRRTDSIGRPLPDRTNVLTQRFPNQGIIITASIOEVLHFGGDDROI
PIIGGAEIYQOQLPFCQRLYTEIQDDFGADTFEPYDRDNREISREMHQATDSGE
YHFVLDKQKQPNVAGCSTDLs"
complement (2059. .2853)
/gene="thyB"
/locus_tag="NE0568"
complement (2059. .2853)
/gene="thyB"
/locus_tag="NE0568"
complement (2059. .2853)
/BC_number="2.1.1.45"
/function="InterPro IPR000398 COG5 COG0207"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="Thymidylate synthase"
/protein_id="CAD84479.1"
/db_xref="GI:30180177"
/translation="MTQYLDLMAHNYQYGHKKSDRGTGTLVFGYGMPPDLOTGPEL
VTTRKCHVKSIIHELLMPLRGETNIDYLKNGSVIWDKADENGDIAPITGHQWRSIA
ASDGTVIDIQVYQIQIKSTPDSRRMIVSAMVNGDLKKKLACVHLPFYVADGRIS
COLYORSADIFLGVPENIASYSLITIMIAOCCDLQGFEPVHTGDAHLVNLHLEQARL
OLEREERAPLPMQINSTVNNIFDGYEDFLHDYDPYPIKAPVAV"
2952. 4454
/locus_tag="NE0569"
2952. 4454
/locus_tag="NE0569"
2952. 4454
/locus_tag="NE0569"
/BC_number="3.4.11.1"
/function="InterPro IPR000819:IPR001589:IPR001687 COG5
COG0260"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="Cytosol aminopeptidase"
/protein_id="CAD84480.1"
/db_xref="GI:30180178"
/translation="WLATLLIONNTPESTRSHLLIIPKTDKIPKGYOLPAEQVIN
DLQROGI PVNRLAETPLSASIQHGLVAVWVDPGSLSEQOTSLKNAQOLNENP
AEIHLAVYGDAQRHRLAEALVAVWVNGAILLVKQKRDDEPLERI CLHGLVDSTG
FAMQCOAQBENFLARGTLIPNELLFRYTRQRIKLAASENQYEBEYSRLAEICA
GARSAYAGSLTDDAALVHLRRTAAGHETGKTIALVKGK:CFPDGHNLPARMYGM
HEDNNGSAVALGVTLAATRADLPINIDCWALIQNHTSPAPKQNDVITLNGWTIETI
VHTDABGRVLADTLTLAAREKPDLLIVATLIGSLITAGSRISGYFTTRDIYAQNA
VAVGRSSGERVCVLPVDSYEDALDSKIDIKOCTTESADHILAACFLHFVNDTPW
LHMDLSACRHQDGLGAVGTETGVTGVTWMSHFLQELIKILQKMSKSRTP"
4505. 4855
/locus_tag="NE0570"
4505. 4855
/locus_tag="NE0570"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible lipase"
/protein_id="CAD84481.1"
/db_xref="GI:30180179"
/translation="WNDEVKQRLQSTETWIRVLFMLFMPIQGSVKFLIVLLAFOLG
STVLTQANRLIKLGRQLAMVYQISLFLFNSQRPFPSSWPSDTONRTSDNADN
RTIPNENKTSWFG"
4559. 4627
/locus_tag="NE0570"
/note="1 probable transmembrane helix predicted by
TMHMM2.0"
complement (4977. .5738)
/locus_tag="NE0571"
complement (4977. .5738)
/locus_tag="NE0571"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"

Db	171851	GGTCCGCTGTCCCGGATACCGCCCGGCTTTTGGCGCGCATATCATGCGCAACCTCA	171911
QY	133	AGCCTCTATCACTACACATGCTGCTGTGTGATACCGGAGGCGATATTGAAATCTT	192
Db	171911	AGCCGCGCATCGTGCACCTGAAACAGGACGTGCTTATCGTCCGATTTCTGCTCTT	171970
QY	193	TACCGTTGA-----TAGTATTTCACCTCTTATGACACTTTATTTGACAGATACAA	246
Db	171971	TACGGCTGACCTGCGCGGCGGTTTCAGCGCTGATGACGCTATACCTGACGGAATGCA	172030
QY	247	CCAGTCCATAGGAATCAAACTAGACAGAGAGACCAGTCCGATTTGGGGTGAAGTTG	306
Db	172031	CATCACCGGAAGAAATTACTCGGCTCAGGCAGTGTATTCGTGCAAGATATCAAGCTT	172090
QY	307	ACCCGCGGGGCGCACCAAAATTTCTAAGATGAGTGAATCTTTTCGGGAAGTGT	366
Db	172091	ACCTTGGTGAAGGACACACATTTCCGATGCGGATGTGACGGAATTT---GCACGCTGCG	172147
QY	367	TACCAATTCTACAGAAATGTTGAGCATATATGCTCGCGATTCATAGAGAGTTA	426
Db	172148	AGGCTACGCTGAGAAAGATGGAAGAGCTGCACATGCTTTACTCTGTGATGGAAGTGG	172207
QY	427	CTATCCTGAGGTGACATGTTTGATAGAGAAAGATATTCATGAAACGGTCTTAAGC	486
Db	172208	TCGATCCTGAGGTGATGTTTTCGACCGGGAATAATTTTCATTGACCGGGTACTGACC	172267
QY	487	CGTTGTGCAAAATTTTCAACAATTAAGTGTGATGAGATGTTTACACACTTTGATG	546
Db	172268	CGCTACTGCAAGCTTTTCCCGGCTGCGATGATATTCGAACATATCATCTTACAGAGAG	172327
QY	547	CTGTTAAGTTTGTGAAATCTTGACTGAAAGATTGTTGCAACCACTGTCAACCCCAAC	606
Db	172328	CAGTGCAGTTTGTGCGAC--GCTCCAAACCGGATGCTGCCACGATTAACCGCACCC	172384
QY	607	ATCTTGTTTTGAACAGAAATCTCTCTTTCAGAGGGGCTTCAACCGATTAATCTGCC	666
Db	172385	ACCTGATGTGAACCGGAATGCCCTGTTTACCGCGGATGCTGTCGCAATCATTACTGT	172444
QY	667	TTCCAGTCTCAAAAGAGATCCACAGGAGGCACTTGTGACCTGTAAACAGTGGAA	726
Db	172445	TGCGGTACTGAAACGGGAAATCATCTCAGGCACTCTGGAAGCACGACAGCGGTC	172504
QY	727	GTAAAGATTTTCTTGGGACTGATGATGCTCCATGATAGACGAAGAAAGTGT	786
Db	172505	ATTCAGATTTCTTCTGTGACCAAGCGCACACATCCGGTCAGAGACAAAGATCAG	172564
QY	787	CTTGTGATGTGCTGTGATTTTACAATGACCTGTAGCCTTGTCAATATGCAAGGTGT	846
Db	172565	CTGTGGGTGTGCGGGTATTTTACAGTCTCATGACCATGAAATTTATGCGAGGTAT	172624
QY	847	TTGAAAGAAATATGACCTGCACAGAGCTGAAGATTCATAGTTTCAATGACACGATT	906
Db	172625	TCGAACAGGCTGCGCGGTGCATGCACTGAGGCTTTTACACAGCTTTCACGCGCCGATT	172684
QY	907	TTTATGGCTTCTTAGAACAATCAAAAGTTAAGTTAGATGAGACCCGATGAGAGTAC	966
Db	172685	TCATATGTTTACACGATATACCATCGGATTTCCGTATCAGGAAAGCTGGAGATTTC	172744
QY	967	CCGATCCTTTCTTATGACATCAGGAGATTTATCCCATGTTTGTGCTGGAATGCTCG	1026
Db	172745	CGAGACAGGTGCAATTTGGGTGAAGAAATTTGATCCCTTCAGAGCAGGGGAACAAGCTC	172804
QY	1027	ACTGTGTCGGGC 1039	
Db	172805	GCTGGAACTGGC 172817	

RESULT 9	AP005085/c	LOCUS	AP005085	303450 bp	DNA	linear	BCT 07-APR-2004
DEFINITION			Vibrio parahaemolyticus DNA, chromosome 2,			complete sequence, 2/6.	
ACCESSION	AP005085	BA000032					

VERSION	AP005085.1	GI:28808465
KEYWORDS	Vibrio parahaemolyticus	
SOURCE	Vibrio parahaemolyticus	
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.	
REFERENCE	1	
AUTHORS	Nasu, H., Iida, T., Sugahara, T., Yamaichi, Y., Park, K. S., Yokoyama, K., Makino, K., Shinagawa, H. and Honda, T.	
TITLE	A filamentous phage associated with recent pandemic Vibrio parahaemolyticus O3:K6 strains	
JOURNAL	J. Clin. Microbiol. 38 (6), 2156-2161 (2000)	
MEDLINE	20295086	
PUBMED	10834969	
REFERENCE	2	
AUTHORS	Makino, K., Oshima, K., Kurokawa, K., Yokoyama, K., Uda, T., Tagomori, K., Iijima, Y., Najima, M., Nakano, M., Yamashita, A., Kubota, Y., Kimura, S., Yasunaga, T., Honda, T., Shinagawa, H., Hattori, M. and Iida, T.	
TITLE	Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism distinct from that of V. cholerae	
JOURNAL	Lancet 361 (9359), 743-749 (2003)	
MEDLINE	22508454	
PUBMED	12620739	
REFERENCE	3	
AUTHORS	(bases 1 to 303450)	
TITLE	Hattori, M., Yamashita, A., Oshima, K. and Shiba, T.	
JOURNAL	Submitted (09-APR-2002) Masahira Hattori, Kitasato Institute for Life Sciences, Kitasato University, Kitasato 1-15-1, Sagamihara, Kanagawa 228-8555, Japan	
MEDLINE	(E-mail:hattori@genome.1s.kitasato-u.ac.jp, Tel:81-42-778-8194, URL:http://genome.1s.kitasato-u.ac.jp/, Fax:81-42-778-8193)	
PUBMED	16-mal:hattori@genome.1s.kitasato-u.ac.jp	
REFERENCE	4	
AUTHORS	This work was done in collaboration with Kozo Makino, Ken Kurokawa, Katsushi Yokoyama, Takayuki Uda, Kenichi Tagomori, Masatomo Najima, Masayuki Nakano, Yoshino Kubota, Shigenobu Kimura, Teruo Yasunaga, Takeshi Honda, Hideo Shinagawa, Tetsuya Iida (Osaka University), Yoshio Iijima (Kobe Institute of Health), and supported by the Research for the Future Program of the Japan Society for the Promotion of Science. This clone was isolated from a patient presenting with acute gastroenteritis.	
JOURNAL	Location/Qualifiers	
FEATURES	1. .303450	
SOURCE	/organism="Vibrio parahaemolyticus"	
	/mol_type="genomic DNA"	
	/strain="O3:K6"	
	/sub_strain="RIMD 2210633"	
	/db_xref="taxon:670"	
	/chromosome="2"	
gene	81. .1574	
CDS	/gene="VPA0300"	
	81. .1574	
	/gene="VPA0300"	
	/note="similar to EMB:CA646187.1 (A1591787) percent identity 68 in 487 aa"	
	/codon_start=1	
	/evidence=not experimental	
	/transl_table=11	
	/product="hypothetical signal peptide protein"	
	/protein_id="BAC61643.1"	
	/db_xref="GI:28808466"	
	/translation="MNLKAVPKALVLTLLASLPISVAVYAAPNANSSSLVQOQMEGNWPTONEAKLVQELYYORALISAYQOMQANVIGMDSESKFGYVVLPIWDRMD SRTLVPENADVVVSMNLDLKEKGPLVVAAPVKYIGWTFDFORTITDVLIGPDA RCGVLLIPNVYDVPVPGKGYPTFTSSTNNVLPRTYWKGGENGADSDAKIAEOR IYPLMDEKNIQPMKFPNAGSORINMYPPTPSVWEKKQGVDEEPIISAISEPRGVU ASIGIVKGFNPVVEEKQLEKAVKTPKMTIQRLGKDEKNTLYKDRYRANA GATSEYQESYLDIDQRAFPQYVSSAPAMWMTMAGSKYPTVSDSKDILLGSH QYKLIHPANPPAKLFWATLVNVVDTGMPETSOILPSKMGPDNIKNTDGSIDLYFGP QLPNGAPESNYIKTIPGRDPLTAIRLVGTGIEFPDQTKPDDVLKIK"	
gene	complement (1835. .2314)	
CDS	/gene="VPA0301"	
	complement (1835. .2314)	
	/gene="VPA0301"	
	/note="similar to DBJ:BAB43758.1 (AP003138) percent identity 57 in 158 aa"	
	/codon_start=1	
	/evidence=not experimental	
	/transl_table=11	
	/product="putative peptide methionine sulfoxide reductase"	
	/protein_id="BAC61644.1"	
	/db_xref="GI:28809467"	
	/translation="MMKITYPAGCICMGVEPMRHLPGVISTEAGRANKTNDTQSEY DQYACVARTPEPDSQVSIETLMGYPFEIIPYISINKGDDVGEKRTGYVSDTKALE IAKSFATRYDDKIAIVEVLPKNVYPSDEHQRILTRPPNDYCHPLDLHKRYKK"	
gene	2493. .2882	
CDS	/gene="VPA0302"	
	2493. .2882	
	/gene="VPA0302"	
	/note="similar to GB:AA64344.1 (L16865) percent identity 52 in 129 aa"	
	/codon_start=1	
	/evidence=not experimental	
	/transl_table=11	
	/product="hypothetical protein"	
	/protein_id="BAC61645.1"	
	/db_xref="GI:28808468"	
	/translation="MKIEHVAILWTERLEJLKGFEYKFNVAVSNDKYHNPKQFSSYFL SRSRGARLEIMGECVTCNCNSHQVTLAHFAFALGSEQAIVDQITKTLVEDGYORI DGPRTYGDGYSCVLDPPGNRIELTV"	
gene	2974. .3714	
CDS	/gene="VPA0303"	
	2974. .3714	
	/gene="VPA0303"	
	/note="similar to GB:AAL20622.1 (AB008775) percent identity 54 in 243 aa"	
	/codon_start=1	
	/evidence=not experimental	
	/transl_table=11	
	/product="putative regulatory protein"	
	/protein_id="BAC61646.1"	
	/db_xref="GI:28808469"	
	/translation="MNPQNEIILQVNDKRKVYVTELSIDIGVSGVTIRODNFLBEO GYIKRHHGATALQSDIDIRLEVRFDIKQLANKAADVAPEVTVLIGSSANALIA RTLAERGDVITITPSAYIAHLIRNTSANITILGCVAOHQESLVGPITKLCINIHFS TAPLIGIDGPHQDGTFSRDMRADIAEAILAKKRRNIVLTDSSKFGQIYSSIKTIVE ISVLTDBKAPKSDLEQLKKGVEVVLG"	
gene	complement (3733. .4089)	
CDS	/gene="VPA0304"	
	complement (3733. .4089)	
	/gene="VPA0304"	
	/note="similar to GB:AAF96321.1 (AB004376) percent identity 76 in 117 aa"	
	/codon_start=1	
	/evidence=not experimental	
	/transl_table=11	
	/product="conserved hypothetical protein"	
	/protein_id="BAC61647.1"	
	/db_xref="GI:28808470"	
	/translation="MHQHEKLVNVERGTNIGATKAFEFHFVGMQFVDYGPDYAASG QGDDGFGYSAEQVSQTSNGALLVFYSSDIHATLEKAVKGGQIIRIFEPGGRPH FTPESGNEFPAWSESH"	
gene	4319. .5860	
CDS	/gene="VPA0305"	
	4319. .5860	
	/gene="VPA0305"	
	/note="similar to REF:NP_231225.1 (NC_002505) percent identity 83 in 510 aa"	
	/codon_start=1	
	/evidence=not experimental	
	/transl_table=11	
	/product="catalase"	
	/protein_id="BAC61648.1"	
	/db_xref="GI:28808471"	
	/translation="MVSTHWRATQMSKSFLLITVGLASTSLQAOITLRDNGAFVGN QNSTIAGNSVLLQDVHLIOKQRFARERIRPERVVARGTGAHGEFVAGDFSLTV"	

gene	5962. .6429	/gene="VPA0306"	5962. .6429	/note="similar to GB:AAF94738.1 (AE004235) percent identity 63 in 152 aa"	percent
CDS	5962. .6429	/gene="VPA0306"	5962. .6429	/codon_start=1	
		/evidence=not experimental		/transl_table=11	
		/product="Amx protein"		/protein_id="BAC61649.1"	
		/db_xref="GI:28808472"		/translation="MKNVPLLSVLFPSIGLSPSLPASQDKLEFYQSYVELFPDAARI	
		GNNEVETFEVSGFPIDQRMNESYTLAWVAAYQGNKDTYRLLESQANNAQLQDRNT		ALMGALIKREISIALNDLNOECSPDLRNKAQGLDKTFAETIYGSEVLKSLQNK"	
		complement(6548. .7201)		/gene="VPA0307"	
		complement(6548. .7201)		/gene="VPA0307"	
		/note="similar to GB:AAF42465.1 (AE002563) percent identity 47 in 200 aa"		/codon_start=1	
		/evidence=not experimental		/transl_table=11	
		/product="putative pyrazinamidase/nicotinamidase"		/protein_id="BAC61650.1"	
		/db_xref="GI:28808473"		/translation="MTYRVANYQTTAVIDVDPKGFSESLCPDELPVAGALDIYPELLKN	
		HAKGSLIKVLSRDLHPKRAMDAETPAMLEPVGLPVIADYKMRHCVLGTGGVELDLGL		PVLLYDPOVKNQMDPDHAPVGIPEHDVADTKTTGAAEFKCKKIDTVVVGGLADLFC	
		VKSSWQALDLSGFYIYNLAATRAVLPDVIIVSAIEMKEKGLFVNADDIIVERFA"		complement(7234. .7959)	
		/gene="VPA0308"		complement(7234. .7959)	
		/gene="VPA0308"		/note="similar to GB:AAF96011.1 (AE004352) percent identity 60 in 241 aa"	
		/codon_start=1		/evidence=not experimental	
		/transl_table=11		/product="conserved hypothetical protein"	
		/protein_id="BAC61651.1"		/db_xref="GI:28808474"	
gene	22.9%; Score 290.6; DB 1; Length 303450;				
CDS	Best Local Similarity 57.4%; Pred. No. 3.3e-63;				
	Matches 584; Conservative 0; Mismatches 424; Indels 9; Gaps 3;				
15	CTCTCAATACACAACCTGATGATTGGACATCTTCATCTCCGATGATGATGATCTTTAAG	74			
104344	CTGACTATTACTCGTCTGACGACTGGACGCTTCACTTACGGAGAGGAGGAGGTTCTAAAA	104285			
75	GCACTTGCTCTGACAGTGCATCATCTTTGGAGAGGCAATAGTCATCCAAATTTGAG	134			
104284	GACACCGTTGGCGATATACAGCCGCTTCAATATGCCCGCGCGCTGATCATCTCCAAACCGTCT	104225			
135	CTCTCTATCACTACCATGCTGCTGTGTGAGGATACCGGAGAGGCGGATTTGAAATCTTTA	194			
104324	CCACCTGTTACCAATGCCAATATGGACATTCGCTTACCCGATTCGAAATTCGAAAAGACAG	104165			
195	CCGTGATAGTATTTCAACCTCTTATGACATTTTATTTGACAGATACACCAAGTCTT	254			
104164	CACGGCGA--ACAATTCGAACCTCTGATGGCGCTTTACTTAACGATTAACACACAGCCA	104108			
255	ATGGAATCAAACTACAGAGAGCCAGGTCGATTTGGGGGTGAAGTTGTAACCTTGCT	314			

Accession	Source	Organism	Reference
DB	104107	GAAGGATTCGCCGACGCAAAAGCAACCGGTTAAATGTCTGACGCAAGCTTTACCCAGCA	104048
QY	315	GGTGCACGACAAATTTCTCAAGATGAGTGACTGATCTTTTCGGGAAGTGTATACCAAGTT	374
DB	104047	GGTGACAGACTAACTCGAGATCTTGCGGTGACCGATCCG--AAGAACATTTACCAAGTT	103991
QY	375	CTACAGAAATGTGTTGAGCATTAATATAGTCCCTCTGCTGTTCAATGAGAGGTTACTAATCC	434
DB	103990	CTCGAAGCANTGAAAGAAAGTCGGATGTCTTTGGTTCATGAGAAAGTCTACCCATCAC	103931
QY	435	GAGGTTGACATGTTTGTATAGAGAAAGGATTTGATTAAGACGGTGTCTTAAGACGGTGTG	494
DB	103930	CACGTTATATCTTCGACCGAGAAAAAGAAATTCCTGATACCGTCTCGCTCTATCGTT	103871
QY	495	CAGAAATTTCCACAAATTTGAAGGTCGAGTGAAGGACATGTTACCAACCTTATGCTGTTAAG	554
DB	103870	AATGACTTCCCAATCTTAAATTCGTTCTTAGAACATATACACACCGCCACACACACACG	103811
QY	555	TTTGTTGAATCTTGCACTGAGAGATTTGTTGACAACTGTACCCCAACATCTGTT	614
DB	103810	TTGTTTAACACCGCTCTGA--TATGTTGCGCCGCTACATTAATGCTGCTACACACTTCTT	103754
QY	615	TTGACAGCAATTCCTCTTCCAAAGGCGCTTACAAACCGCAATTAATTAATGCTTCAAGTC	674
DB	103753	TTCAACCGCAATCAATGCTTGTGCGTGAATTAAGCTTCAATTTCTACTGCTTACCGCAAT	103694
QY	675	CTCAAAAGAGATTCACAGGAGCACTTGTGTAGCTGTATACACAGTGAAGAGTAAAGA	734
DB	103693	CTAAAGCGCAACCCACACAGCAAGCGCTAATTCGAAGCCGCAACACAGTGTAGCAAGAG	103634
QY	735	TTTTTTCTTGAGAGTATATGATGCTCTCATATATAGACGAAGAAAGTGTCTTTGGA	794
DB	103633	TTTTTCTTGATGACGACTGTCTGCTCGATGCAAAAGCGGAAAGATCGGCTTGGCGT	103574
QY	795	TGCGCGGATTTAAATGACACTGTAGCCTTTGTCAATATGCGAAAGTGTGGAAG	854
DB	103573	TGTGCGGTTCTTAACTAGTCCGACGCGCGGTAGAACCTTATGCAAGAGCTTTGAAAA	103514
QY	855	GAAGATGACCTGACAGACCTTGAAGCAATTCATGCTTCAATGACAGCAAGTTTTTAAGG	914
DB	103513	GAAGGCAAGCTGAAAACTTGAAAGCGTTGSCAAAGCTTCAACGCTCTGAATTTTACGGT	103454
QY	915	CTTCTAGAACAACTCAAGATTAATGATGATGACGSCCATGGAAGGTATCCGATCC	974
DB	103453	ATTGGCGCAACGCAATGACGTGACGCTGAAAAAATCGGCTTGGGATGTTCTGATATC	103394
QY	975	TTTTTCTTATGATGAGAGATATTAATCCCATGTGTTGCTGTGTAATGCTCGAGCTG	10331
DB	103393	ATGCCATTTGGCAACGACATCGTTGTGCTTAATTCGTGCAACGAACAAAATGAGTGG	103337
RESULT 10	AE016925/c		
LOCUS	208524 bp	DNA	linear
DEFINITION	Chromobacterium violaceum ATCC 12472		BCT 14-OCT-2003
ACCESSION	AE016925		complete genome.
VERSION	AE016925.1		GI:34105514
KEYWORDS	Chromobacterium violaceum ATCC 12472		
SOURCE	Chromobacterium violaceum ATCC 12472		
ORGANISM	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;		
REFERENCE	1 (bases 1 to 208524)		
AUTHORS	Brazilian National Genome Project Consortium		
CONSTR	The complete genome sequence of Chromobacterium violaceum reveals		
TITLE	remarkable and exploitable bacterial adaptability		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11660-11665 (2003)		
MEDLINE	22882880		
PUBMED	14500782		
REFERENCE	2 (bases 1 to 208524)		

AUTHORS	
Vasconcelos, A.T.R., de Almeida, D.F., Almeida, F.C., de Almeida, L.G.P., de Almeida, R., Goncalves, J.A.A., Andrade, E.M., Antonio, R.V., Aratipe, J., de Araujo, M.F.F., Filho, S.A., Azevedo, V., Batista, A.U., Batista, L.A.M., Batista, D.S., Belo, A., Vander Berg, C., Blamey, J., Bogo, M., Bonato, S., Bordignon, J., Brito, C.A., Brocchi, M., Butty, H.A., Camargo, A.A., Cardoso, D.D.P., Carneiro, N.P., Carraro, D.M., Carvalho, C.M.B., Cascardo, J.C.M., Cavada, B.S., Chneire, L.M.O., Pasa, T.B.C., Duran, N., Fagundes, N., Falcão, C.L., Fantinatti, F., Farias, I.P., Felipe, M.S.S., Ferrari, L.P., Ferro, J.A., Ferro, M.I.T., Franco, G.R., Freitas, N.S.A., Furlan, L.R., Gazzinelli, R.T., Gomes, E.A., Goncalves, P.R., Grangeiro, T.B., Grattapaglia, D., Grisard, E.C., Guimaraes, C.T., Hanna, E.S., Hungria, M., Urdim, S.N., Laurino, J., Leol, L.C.T., Passarella, L., Lima, A., Loureiro, M.F., Lyra, M.C.P., Macedo, M., Madeira, H.M.F., Manlio, G.P., Maranhão, A.Q., Martins, W.S., de Mauro, S.M.Z., de Medeiros, S.R.B., Meisner, R.V., Menck, C.F.M., Moreira, M.A.M., Nascimento, F.F., Nicolas, M.F., Oliveira, J.G., Oliveira, S.C., Paixão, R.F.C., Parente, J.A., Pedrosa, F.O., Pena, S.D., Pereira, J.O., Perreira, M., Pinto, L.S.R.C., Pinto, L.S., Porto, J.I.R., Potrich, D.P., Neto, C.E.R., Reis, A.M.M., Rigo, L.U., Rondinelli, B., dos Santos, E.B.P., Santos, F.R., Schneider, M.P.C., Seunarez, H.N., Silva, A.M.R., da Silva, A.L.C., Silva, D.W., Silva, R., Simoes, I.C., Simon, D., Soares, C.M.A., Soares, R.B.A., Souza, E.M., Souza, K.R.L., Souza, R.C., Stefens, M.B.R., Steindel, M., Teixeira, S.R., Umenyi, T., Vettore, A., Wasseem, R., Zaha, A. and Simpson, A.J.G.	
TITLE	
Submitted (22-JUN-2003) Labinfo, LMCC - Laboratorio Nacional de Computacao Cientifica, Rua Getulio Vargas 333, Petropolis, RJ 26651070, Brazil	
JOURNAL	
FEATURES	
source	Location/Qualifiers
gene	1. .208524
gene	/organism="Chromobacterium violaceum ATCC 12472"
gene	/mol_type="genomic DNA"
gene	/strain="ATCC 12472"
gene	/db_xref="taxon:243365"
gene	/complement(96. .563)
gene	/gene="p1B2"
gene	/locus_tag="CV4209"
gene	/complement(96. .563)
gene	/gene="p1B2"
gene	/locus_tag="CV4209"
gene	/note="identified by sequence similarity; putative; ORF located using Genemark/Blastx/COG2165"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="type IV pilin"
gene	/protein_id="AA061869.1"
gene	/db_xref="GI:34105515"
gene	/translation="MKQRKQGFLLIELMIVAIVGILAAIALPAYODYTKARVSEGLALADAKTAVTEYATNNAMASGTAPFNTTGLAASITNSVSSINVLANGVQA VNSTVNGALVDIVPTVSAGSVQVCTYTGSGTAAIASQLSANWVSTCRQ"
gene	838. .2184
gene	/gene="b10A"
gene	/locus_tag="CV4210"
gene	838. .2184
gene	/gene="b10A"
gene	/locus_tag="CV4210"
gene	/EC_number="2.6.1.62"
gene	/note="identified by sequence similarity; putative; ORF located using Genemark/Blastx/COG0161"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="adenosylmethionine-8-amino-7-oxononanoate transaminase"
gene	/protein_id="AA061870.1"
gene	/db_xref="GI:34105516"
gene	/translation="MSNQAWLERSFPAVAVHCTQMKRHERLPIVPIASADGWLITDPD GSKRYLDGVSSWVNLFGHGRPIKQAVKQDSEHMLAGFTRPVVELSERLAALS GLGHAFYSGDASATEIALKMSFYMKVGRPOKTRFVSLNSYHGEVGLAATDVP LFSSTYATLKRGLRAPSADAROVGERADVALRAARSLNLLAKGSGIAIVYF PLVQAGAMAHDPYLTETLRICDQYVHILADEIVAGRTGTSFPYQOAGITPDP LCLSKGITGFLPLSCVLTREIITQAFTHDVARGFLSHSYTGNAALAAVLVDI
gene	2347. .3816
gene	/locus_tag="CV4211"
gene	2347. .3816
gene	/locus_tag="CV4211"
gene	/note="identified by sequence similarity; putative; ORF located using Glimmer/Blastx"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="conserved hypothetical protein"
gene	/protein_id="AA061871.1"
gene	/db_xref="GI:34105517"
gene	/translation="MRQARPGVVPAGPRPDGSRALQAQALSBGAGPLSAAAGERS PARRGRLLAQNNAAGHQAQVAVRPGVPCQAPARKLLPPAAGARTTAQGAAYHDD GGAALPEHOPAKULASLNLGCLPPVAVAGAAARLACGAVARRRRRACGLPAAAR SRHOPAVARRRPAIGATLRKCATDVAAGCGDGPGLCGPERGCPALISACQ IAVPARLLPARGTRAGRDETQPTTAQDGRGRRLSGARRSPAAEPSSGPMASAPV TQACPGGRDGGDHRRRRGLAPRRQQLAARRRRRREORPAPARLLPADCRQPE RIMPARAPSAALARRRGAAGPARPAGRVVYSALGIDPAALRQSGGMRRDAGP AAAAGGHAQHHPQRQPFPTGLAARQRPAGAAVAPRSLQPVANVOAGRRRATG PRQAASALPPFPADGIPQQRALITWRKRASFVTLPSL"
gene	3845. .4963
gene	/locus_tag="CV4212"
gene	3845. .4963
gene	/gene="prob"
gene	/locus_tag="CV4212"
gene	/EC_number="2.7.2.11"
gene	/note="identified by sequence similarity; putative; ORF located using Glimmer/Genemark/Blastx/COG0263"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="glutamate 5-kinase"
gene	/protein_id="AA061872.1"
gene	/db_xref="GI:34105518"
gene	/translation="MHSYTHASDRIVVKVGSLLVTNDRGLDINALARWAEVAELKR RKOYVIVSSGAIABGCRLGWTRPKGHEILQAAAAGVGMCGOAYESAFRFGART AQTLLHEDLADRTYRLNARSTLTSLNANVYIINENDVATSEIRFGNDTLGALV TNLIEDALVITLTDQRLYSADPRKHDFEFTIQAERGDERLEDMAGGASSVGTGM ITKILAKRAARSGAATVIASGRPEPLSLADGEAIGTQLVATNRMAKRWLADH LKLAGRLDDGALAALIRERGTSLPLPGVSAVGDPLRGEAVACVDAHEVARGLVN YSSDEARQIMRKSRTREIHALGYIVEPELHRNMVAL"
gene	5057. .6130
gene	/gene="pofF1"
gene	/locus_tag="CV4213"
gene	5057. .6130
gene	/gene="pofF1"
gene	/locus_tag="CV4213"
gene	/note="identified by sequence similarity; putative; ORF located using Glimmer/Genemark/Blastx/COG0687/TC:3.A.1.11.2"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="polyamine ABC transporter system, substrate-binding protein"
gene	/protein_id="AA061873.1"
gene	/db_xref="GI:34105519"
gene	/translation="MNKIALAALLAAVACGSAAEVDLHIYNNMNAISPDTAKPFGQSC KCTLVQDYGDNEEMLAKLAAGAKGYDMVPTFAVNTLLKQGLQPLDRLPNWN LNSGYALNPPQGNRYAAPTVASLTLLGYNAITOLKKAEGVSGNSMALVFDALLA KIKGKTYVDSQSEIEMAAALMYIKDQANSNPNPADWKAARETIRAKRYMAFNNQSYI KEITVGINIWAIGISNDLFOAQDPAKRAKPPPELAYRPOEGVTLAIDNNTLLKIDAR PDLAKHINFMIDGKANSIENQIGATNPVKAABAFKPOIKANPVIMLDPSSKQIYA LKDLVDYSREIRLMSQVXIKR"
gene	complement(61173. .6757)
gene	/locus_tag="CV4214"
gene	complement(61173. .6757)
gene	/locus_tag="CV4214"
gene	/note="identified by sequence similarity; putative; ORF located using Glimmer/Genemark/Blastx/COG3575"
gene	/codon_start=1
gene	/transl_table=1

QY 435 GAGGTGACATGTTTGTAGAGAAAAGTATTCATTGAAACGGTCTAAGACCGTTGGTG 494
DB 662 GAAGTGCACATTTTGTGACCGTGAAAAGACCTTCTTACACACCGTGTGCGCGATGIG 721
QY 495 CAGAAATTTTCACATGTAAGGTGATGATGAGACATGTTACACATTTGATGCTGTTAAG 554
DB 722 AATGATTTCCCGACACTGAAATTTGTCTTGAACATCAACACCGCGCATGACATGACC 781
QY 555 TTTGTTGAATCTTGACATGAGGATTTGTTGACGAACTGTCAACCCCAACATCTTTGTT 614
DB 782 TTTGTGCAA---CAAGCAGCGGATTAACGTTGCGGCAACGATTAACCGGCAACATTTGCTG 838
QY 615 TTGAAACAGGAATTTCTCTTCCAAAGGGGCTTACACCGCATATTTACCTGCTTCCAGTC 674
DB 839 TTTAAACCGTAAACACATGCTGTTGGGATTCGCCACATTTTCTACTGTTTGCATATC 898
QY 675 CTGAAAGAGAGATTCACAGGAGGACCTGTGTGACCTGTATACAGTGAAGTAAAGA 734
DB 899 CTAAAGGCTGACCTCAACACACGCGTTAGTGGACAGGCAACCTTGGCAGCAAGAA 958
QY 735 TTTTCTTGGGACTGATAGTCTCTCATGATAGACGAAAGAAAGTCTTCTTGTGA 794
DB 959 TTCTTCTAGGTACGACCTGCGCCGACGCAAGGCGCAAGAGCGCTTGTGTGC 1018
QY 795 TGTGCTGATTTTACATGACCTGTGACCTTGTCAATATGCGAAGGTGTTGAAG 854
DB 1019 TGGCAGAGTCTTACACAGCCCATGACGCTCTTGAATGATGCGAAGTGTGAAAA 1078
QY 855 GAAATGACACTGCAAGCTTGAAGCACTGACCTTCAATGAGACCATTTTATGGG 914
DB 1079 GAAGTAACTAGAAATCTGGAAGCGTTGCTTACTTAATGACCTGATTTCTACGCG 1138
QY 915 CTCTCTAGGAACTCAAGATTAAGTTAGTTAGTAAGACCGCATGAAAGTACCGAAATCC 974
DB 1139 CTACTCTGCAATCAAGAGACCGCTCACGCTCAACGAAGCGCTGCGCATGTTCTACGCG 1198
QY 975 TTTTCTTATGATCAGAGATATTAATTCCTCATGTTGTGTGAAATGCTGACTGG 1031
DB 1199 ATGCCGTTTGGCAGCATATCGTGTGCAATCCGCGCGGTGAATAATGAGTGG 1255

RESULT 13
BX572091 349082 bp DNA linear BCT 10-0UN-2004
LOCUS Prochlorococcus marinus MED4 complete genome; segment 2/5.
DEFINITION BX572091 BX548174
ACCESSION BX572091.1 GI:33639658
VERSION complete genome.
KEYWORDS Prochlorococcus marinus subsp. pastoris str. CMP1986
SOURCE (Prochlorococcus marinus MED4)
ORGANISM Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
Prochlorococcus.
REFERENCE 1 (bases 1 to 349082)
AUTHORS Roca, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P.,
Ahlgren, N.A., Arellano, A., Coleman, M., Hauser, L., Hees, W.R.,
Johnson, Z.I., Land, M., Lindell, D., Post, A.F., Regala, W., Shah, M.,
Shaw, S.L., Steglich, C., Sullivan, M.B., Ting, C.S., Tolonen, A.,
Webb, E.A., Zinser, E.R. and Chisholm, S.W.
TITLE Genome divergence in two Prochlorococcus ecotypes reflects oceanic
niche differentiation
JOURNAL Nature 424 (6952), 1042-1047 (2003)
MEDLINE 22825698
PUBMED 12917642
REFERENCES 2 (bases 1 to 349082)
AUTHORS Larimer, F. and Roca, G.
CONSRTM Prochlorococcus genome consortium
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2003) Submitted on behalf of the Prochlorococcus
genome consortium, the DOE Joint Genome Institute, Production
Genomics Facility, 2800 Mitchell Drive, Walnut Creek, CA 94598,
USA, and the Genome Analysis Group, Oak Ridge National Laboratory,
1060 Commerce Park Drive, Oak Ridge, TN 37831, USA.

FEATURES
source larimerf@ornl.gov
location/Qualifiers
1..349082=Prochlorococcus marinus subsp. pastoris str.
/organism="Prochlorococcus marinus subsp. pastoris str."
/mol_type="genomic DNA"
/CDS="CMP1986"
/strain="MED4"
/db_xref="taxon:59919"
/note="synonym: Prochlorococcus marinus MED4, strain
coidentity: CMP1378 = MED4"
2503..2835
/locus_tag="PMM0369"
2503..2835
/locus_tag="PMM0369"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="possible Ribosomal protein S14p/S29e"
/protein_id="CAE18828.1"
/db_xref="GI:33639659"
/translation="MLQKRAKLETFDRIRERADIVIEGLAFSSITSYVSKYTTIR
RQARRIAVDAYKIVRTDIESDIDRKEMTSKLVCLBENTHLMKKEQYSAVATNARV
LMRLIRLE"
3283..4968
/locus_tag="PMM0370"
3283..4968
/locus_tag="PMM0370"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative cyanate ABC transporter, substrate
binding protein"
/protein_id="CAE18829.1"
/db_xref="GI:33639660"
/translation="MLNPVNSSTKRYNFKCPSADNSDOFQDONSLSNLNLPANN
HGYLCAVCEBPPIGNDHENFMSMPDPVPIISDFSMGLIKQAYVADAISTENR
DLFVKNASNGDPYKEXLEIRELAKAGLQAPFAAPGAPQAGFFATPACASSPGRRA
PLSLAAGALITVACSONKPKPDPIDPPIINNEKOTLRVGFPIICASPIIVA
EPWGFYKRFMDYKVMKPSKGVARDALNGELDAYTMLAMPISMTLIGCTSPFSIX
LASIENTNGAIVVANKRKGKVPADMKQPIFGVPPISMNILLIYILAKGVDPD
KDVOIRVPPPSIAQLVAGDIDAYLMPDPFNQRAVEDAGFIHLTKELMPGHPCCA
FAAGEPWINEPETERALNKSIIIDAAVSTSPSRKEVAKAISGRGPLANDPEVEAV
LTGKFEDGKLTQNVPRIDPKPYMOSFSHMIQSOLIRMDLGAVAIAIGADFNPS
ASIFLIDBAQALREKGNPPTARFREVLAYBFDPSNPIKYKQDKIDGAF"
4999..5781
/locus_tag="PMM0371"
4999..5781
/locus_tag="PMM0371"
/locus_start=1
/evidence=not_experimental
/transl_table=11
/product="putative cyanate ABC transporter"
/protein_id="CAE18830.1"
/db_xref="GI:33639661"
/translation="MNRVISKRLTFVVTIFSLIFITIMOIAGOGLLVKNPFSFKT
INALIMVSDPFPNGNDIGCNLLISLRVLYGSLAVLVPIGLVGMKLAQ
SSLNIPVQLKPVSLMLPGLVFFVNSSEITGTFVITFITSIMPTLINTAGVASINP
DYLRVSKSLGSKSITIRVILPALIMNIIAGMKISMGTAULVVAEMLMGTGIGIF
IMENMNLSENIFVAILIIGFTFIIDQFPGFIQDKFDYSI"
4999..5088
/locus_tag="PMM0371"
/note="Signal predicted by SignalP 2.0 HMM (Signal peptide
probability 0.892) with cleavage site probability 0.817 at
residue 30"
order(5023..5082,5212..5280,5404..5472,5590..5658,
5677..5745)
/locus_tag="PMM0371"
/note="5 probable transmembrane helices predicted by
TMHMM2.0"
5798..6652
/locus_tag="PMM0372"
5798..6652
/locus_tag="PMM0372"
CDS gene misc_feature sig_peptide

```

gene
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative cyanate ABC transporter"
/protein_id="CAE18831.1"
/db_xref="GI:33639662"
/translaction="MORNSKKRSTHLEAKIGKIFFEKKFPDPYAKRHSKFEAFI
KNNIKLEPNTSPVSIICPSGGKSTLNTLAGLQAPTSGKIIIDGLAIKGGPPRGVY
FQNYALMPEUTTTIQTIERPALVTPEKTPDIKERAINIPLMWGEKRSKPEPESG
GKMKORVIALRSLINPNMLIMDEPFGLDALTSYQSEILSIIMRKNVTYVLLTHSI
EHALLSIKKIILMSGSPATISEEISVDLRPRDRNIEQDPKVIKIMRLEKHLRE
TVAVEEVN"
6685..7128
/genes="cyns"
/locus_tag="PMN0373"
6685..7128
/genes="cyns"
/locus_tag="PMN0373"
/EC_number="4.2.1.104"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="Cyanate lyase"
/protein_id="CAE18832.1"
/db_xref="GI:33639663"
/translaction="MTLPSELTKLLKAKKENLSPEDLGNLIRDEWVASLFGYQAT
VSEKEDATSLAALDITSDKEDSTPYKGLDVIPTDPLITRFELIMQVGLPWKD
VIOKRFSDGIMSIDSIEVDKVEDPKDRLVLYKCGKFLPYKKW"
complement (7213..7404)
/locus_tag="PMN0374"
complement (7213..7404)
/locus_tag="PMN0374"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="metA/Hcf106 family"
/protein_id="CAE18833.1"
/db_xref="GI:33639664"
/translaction="WNIFGSLPEIIVILILGLIFGPKRLPOLGKTIGKTLKGITRA
SKFESEIKITLQLEND"
complement (7336..7392)
/locus_tag="PMN0374"
/locus_tag="PMN0374"
/notes="1 probable transmembrane helix predicted by
TMHMM2.0"
7504..7827
/locus_tag="PMN0375"
7504..7827
/locus_tag="PMN0375"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="possible Type II intron maturase"
/protein_id="CAE18834.1"
/db_xref="GI:33639665"
/translaction="MERSMDPYKFFPKEDFNLFEIPLHKEFOKEIIDLGYLQENIN
QFNVPGQFCVMNGVIRKIYHDLPIFLMDKKPRDNINIKMWLDSTINLSTIDIKVRES
SIXI"
complement (7843..8211)
/locus_tag="PMN0376"
complement (7843..8211)
/locus_tag="PMN0376"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical"
/protein_id="CAE18835.1"
/db_xref="GI:33639666"
/translaction="MKLSLSAVLFLFTLSFAQEKLNVTSDSIOISIKGNPAIG
NVIKIRNNNPEASNKLTYYDKDATTGLVGNVFNVLSEGLSIKSGIDSLITFTD
SGLFKRNISGNRVKTKKF"
complement (8152..8211)
/locus_tag="PMN0376"
/notes="Signal predicted by Signalp 2.0 HMM (signal peptide

```

[illegible]

[illegible]

to residues 1 to 340 of 344 from E. coli K12 : B1601;
residues 10 to 349 of 361 are 62.05 pct identical to
(A8008764) putative membrane transport protein [salmonella]
typhimurium LT2"]
/codon_start=1
/transl_table=11
/product="putative transport protein"
/protein_id="AAM85307.1"
/db_xref="GI:21958543"
/translation="MAVRIGQLIMTPVENOGRLVAMLAMLVVMAGIKRASPYYV
PRLATFLAVINLPLVKLERIKIPRLTAVILVTYIIILMTVLFGRLSGSINPASP
LPGYGMGLMEKRDQEPALFNIEFSDVIDMKYVSPSMNPPRRLSHLSGMTST
FLMLTVMLEFEVERMPYKQVLFDPNBEKGNAILKRALNGVTHYIVLTKTISTATG
VLMFLAAMGVAFALMGLAFVLYNIPINGISALAIPIITQALINPQADAFVGTG
VYILMLIGNIIIDPRMGRLGGLSTVYFSLIFMGLIGPVGMLLSVPLTIKIVL
ELTPGYFPAVLIDGKPAKVEQAL"
complement (4368. .4481)
/gene="Y1740"
complement (4368. .4481)
/gene="Y1740"
/note="residues 4 to 37 of 37 are 47.05 pct identical to
residues 12 to 45 of 46 from E. coli K12 : B1058"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAM85308.1"
/db_xref="GI:21958544"
/translation="MFKLSMIVNNVREHLLLYICSLVLGIDIPFFY"
complement (4825. .5466)
/gene="Y1741"
complement (4825. .5466)
/gene="Y1741"
/function="unknown"
/note="residues 10 to 199 of 213 are 28.71 pct identical
to residues 2 to 187 of 199 from GenPept : >gblAAC46458.1
(U21727) entericidin R [Citrobacter freundii]"
/codon_start=1
/transl_table=11
/product="hypothetical"
/protein_id="AAM85309.1"
/db_xref="GI:21958545"
/translation="MLSDQVNMKKILIIDCEFYTSRGVNTYLSVSIKLDVPTVE
EALIALQEQPDIIIVNLQYCRKGHCPLAOPHYCTPAKYIYLDASYPESTPI
PLTBSVSMPIPTKSVSILAKKH.PNLLRRLPGSLERGLDIGNOQASLFSQCEHKM
CYMTMEPYRIARLKINSGTSVYSHKRHTEKIKVNRNLELCFINNPKIY"
complement (5706. .6821)
/gene="sola"
/note="synonym: Y1742"
complement (5706. .6821)
/gene="sola"
/function="putative"
/note="residues 1 to 371 of 371 are 64.24 pct identical to
residues 1 to 372 of 372 from E. coli K12 : B1059"
/codon_start=1
/transl_table=11
/product="arcsine oxidase-like protein"
/protein_id="AAM85310.1"
/db_xref="GI:21958546"
/translation="MDYDLIVIGSGSVGSAAGYVASOGLNVLMTSDMAMPHQASH
GETRIMRAVSEGERKIVPLVRAQALMDLQATEKLFQAGVYINIGPNSITKATVE
ORSADQYDLPEVLTNSTQIREKMPVTPVDNVAIVFEOSGLSELAIVLTKATYE
AGGCIILFNCPTVLAIESHQADVVTIDGYSATKVVNAGVWSEVLPTLPVTPKRV
FSMQADLEVEANHPFAFVEMPDNIIVYGGPANDALKLGKHHGQALIESAQKRP
FGRAEDQTEVFSFLRHPHVGVCIRGACSYDMSDPEDPTIDTLFEDERWVWVSGL
SGHGFADTVLGEVTLAFQDKRSPIDISATSLAKR"
/note="sola"
/repeat_region

```

/function="S and transposon related functions"
/notes="18268, residues 1 to 402 of 402 are 100.00 pct
>BJ04C13227.1 (AF053947) transposase [Yersinia pestis]"
/codon_start=1
/transl_table=1
/product="putative transposase"
/protein_id="AA085311.1"

```

Query Match	21.8%	Score 277.2;	DB 1;	Length 10421;
Best Local Similarity	55.9%	Pred. No. 1.1e-59;		
Matches 569; Conservative	0.0	Matches 443		

QY	15	CTCTCAATCAGACAACCTGATGATTTGGCATCTTTCATCTCCGTAATGTGATGTTCTTAAAG	74
Db	9210	CTAAAAATTGCGCGCCGGAGTGAATGCGCAATTCATCTACGTATATATAAAATGCTCACT	9261
QY	75	GCAGTGTCTCTCAGTGCACATCACTTTGGAGGGCAATAGTCATGCGCAATTGGAAG	134
Db	9270	ACCGGTGGCCCTTATCTCCGAAGTATGCGCCGCGCTATTGTTATGCGCAATCTAGCG	9322
QY	135	CCCTCTATCACTACCACTGCTGCGCGTAAGCATACCGGAGAGCGATATTGAATCTTTA	194
Db	9330	CAGGCATTTGAAACGGTTGCGCATGCTATTGCTTATCGGAGCGTAATTTTACACGGT	9388
QY	195	CCCTGTATAGTGAATTTCAACCTCTTATGACATTTATTTGACAGATACCAACAGTCT	254
Db	9390	CTGCGGGCCATTAATTCACCCCGTTGATGACGTGTACCTGACTAATAGCCTTGATCT	9449
QY	255	ATGAAATCAAACTAGACAGAGAGACCGAGTCGATTTGGGTGAAGTTGACCTGCT	314
Db	9450	AAAGAGTTGACACGGGTTTATGACAGAGCGTTTTCACGGGCGCAACTATCCGGCG	9509
QY	315	GGTGCCACGACCAATTTCTCAAGTGAAGTGAATCTTTTCGGAGAGTGTTCACAGT	374
Db	9510	AATGCAACCAACCACTCCACTCAGCTGATCTGCAT---CCGGGCAATTACCCTGTG	9566
QY	375	CTACAAGAATGTTGAGCATTAATGTCCTCGCGGTGTTATGAGAGGTTACTAATCT	434
Db	9567	TTTGAACAATGCAAAAGATAGGATGCGCTCTTATTCAGGTGAGTAAAGATGCG	9626
QY	435	GAGTTGACATGTTGATGAGAAAAGTATTCATGAAACGGTTCYAGACGGTTGGT	494
Db	9627	GCGGTGACATTTGATCGTGAAGCCCGTTTATGACCAATTTTATGAGCCCATTCGC	9686
QY	495	CAGAAATTTCCACATTTGAAGTGATGAGAGATGTAACCACTGATGCTGTTAAG	554
Db	9687	CAAAAGTTCCGGAATTAATTTGCTCTTGAGCATTAACGACCAAGATGCGGACGAT	9746
QY	555	TTTGTGGAATCTGCACTGAAGGATTTGTTGCAAGCATCTGCACCCCAACAATCTGTT	614
Db	9747	TATGTCGTGAGGCA--ATCGTTTCTTGGGGCAACGTCACGCGCAACAACCTGAATG	9803
QY	615	TTGAACAGGAATCTCTCTTCCAAAGGGGGCTTACAAACCGATTAATTAAGCTTCCAGTC	674
Db	9804	TTTAAACCGCATCAATGCTGTGATGAGGGGATATGCCCCCACTGTTCTGCTCGCAATA	9863
QY	675	CTCAAAAAGAGATCCACAGGAGGCACTTGTTGTCAGCTGTAAACAAGTGAAGTAAAGA	734
Db	9864	TTGAAGCGCAGCAACCATCAGCAAGCATTTGCGGCGAGCGGTGCGCATGTTCTATGCG	9923
QY	735	TTTTTCTTGGGACTGATGATGCTCTCATGATTAACGAAGAAAGATGTTCTTTGGA	794
Db	9924	TTCTTCTTGGGACGATTCAGTCCCAATGCGCAACATCTGAAGAGATCATCTTTCGCGC	9983
QY	795	TGTGCTGTATTTTACATGCACTGTGAGCTTGTGAGATATATGCAAGGTGTTGAAGAAG	854
Db	9984	TGTGCGGGGTATTTCAACGCCCAAGGGGCAATGCTGTATGCTTCCGTGTTGAGGA	10043
QY	855	GAATAATGCACTGCAACAAGCTTGAAGCATTCACATGCTTCAATGACCAAGATTTTATGGG	914
Db	10044	CTGAATGCAATTCGACATCTGGAACGTTTTCGCGCTTAAATGCGCCACGATTTTATGGC	10103

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 12, 2005, 07:04:55 ; Search time 80 Seconds
(without alignments)
1413.190 Million cell updates/sec

Title: US-10-070-277-2

Perfect score: 1812
Sequence: 1 MELSTQPDMDHMLHRLDGDV.....SGDIIMFAGXGMLDWLPAFL 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep:*
- 18: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1528	84.3	345	US-10-451-554-7	Sequence 7, Appl1
2	1458	80.5	359	US-10-437-963-189113	Sequence 189313,
3	1454	80.2	379	US-10-424-599-149491	Sequence 149491,
4	1067.5	58.9	364	US-10-282-122A-48169	Sequence 48169, A
5	1058	58.4	352	US-10-282-122A-50426	Sequence 50426, A
6	1039.5	57.4	350	US-10-282-122A-77674	Sequence 77674, A
7	1008.5	55.7	354	US-10-282-122A-51586	Sequence 51586, A
8	983	54.2	344	US-10-282-122A-65811	Sequence 65811, A
9	979	54.0	344	US-10-282-122A-65238	Sequence 65238, A
10	978	53.8	348	US-10-282-122A-59728	Sequence 59728, A
11	975	53.8	348	US-10-282-122A-66489	Sequence 66489, A
12	971	53.6	348	US-10-282-122A-43220	Sequence 43220, A
13	962	53.1	348	US-10-282-122A-67595	Sequence 67595, A

14	959	52.9	347	US-10-282-122A-69927	Sequence 69927, A
15	956	52.8	348	US-10-282-122A-78404	Sequence 78404, A
16	952.5	52.6	344	US-10-282-122A-44700	Sequence 44700, A
17	949	52.4	348	US-10-282-122A-75878	Sequence 75878, A
18	943	52.0	354	US-10-282-122A-63114	Sequence 63114, A
19	934	51.5	348	US-10-282-122A-72752	Sequence 72752, A
20	837	46.2	260	US-10-767-701-31604	Sequence 31604, A
21	569	31.4	335	US-10-282-122A-54267	Sequence 54267, A
22	541	29.9	330	US-10-282-122A-58744	Sequence 58744, A
23	539.5	29.8	190	US-10-282-122A-55550	Sequence 55550, A
24	527	29.1	339	US-10-335-977-6923	Sequence 6923, Ap
25	527	29.1	371	US-10-335-977-6924	Sequence 6924, Ap
26	420	23.2	270	US-10-767-701-51620	Sequence 51620, A
27	402	22.2	136	US-10-282-122A-68916	Sequence 68916, A
28	365	20.1	87	US-10-424-599-149490	Sequence 149490,
29	358	17.5	86	US-10-767-701-37668	Sequence 37668, A
30	262	14.5	82	US-10-424-599-250183	Sequence 250183,
31	209.5	11.6	146	US-10-335-977-6922	Sequence 6922, Ap
32	144	7.9	396	US-10-282-122A-52409	Sequence 52409, A
33	131.5	7.3	426	US-10-282-122A-60843	Sequence 60843, A
34	128	7.1	424	US-10-282-122A-52160	Sequence 52160, A
35	121.5	6.7	427	US-09-815-242-10858	Sequence 10858, A
36	121.5	6.7	427	US-10-282-122A-57073	Sequence 57073, A
37	119.5	6.6	426	US-10-282-122A-57626	Sequence 57626, A
38	117	6.5	421	US-10-282-122A-52993	Sequence 52993, A
39	115.5	6.4	428	US-10-156-761-14392	Sequence 14392, A
40	111	6.1	2198	US-10-369-493-5496	Sequence 5496, Ap
41	110	6.1	422	US-10-282-122A-72383	Sequence 72383, A
42	108	6.0	422	US-09-815-242-5884	Sequence 5884, Ap
43	107.5	5.9	425	US-10-282-122A-70902	Sequence 70902, A
44	106.5	5.9	422	US-10-472-928-2332	Sequence 2332, Ap
45	105.5	5.8	323	US-10-424-599-155522	Sequence 155522,

ALIGNMENTS

```
RESULT 1
US-10-451-554-7
; Sequence 7, Application US/10451554
; Publication No. US20040111769A1
; GENERAL INFORMATION:
; APPLICANT: Rodolphe Arthur Kanhouou
; APPLICANT: Ramon Serrano Salom
; APPLICANT: Roque Ros Palau
; TITLE OF INVENTION: Sugar Beet genes involved in stress tolerance
; FILE REFERENCE: 1187-24
; CURRENT APPLICATION NUMBER: US/10/451,554
; PRIOR APPLICATION NUMBER: PCT/EP01/15093
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: EP 00870319.1
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/271,656
; PRIOR FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-10-451-554-7
Query Match 84.3%; Score 1528; DB 16; Length 345;
Beet Local Similarity 82.0%; Pred. No. 5.6e-147;
Matches 282; Conservative 34; Mismatches 28; Indels 0; Gaps 0;
Cy 1 MELSTQPDMDHMLHRLDGDVYKAVSHHFGAIVMPNKKPITTTAAVAAREATLK 60
Db 1 MELTLTPDDMDHMLHRLDGDVLAAYVAHSAHFGAIVMPNKKPITTTGAIAVRKSTME 60
Cy 61 SLVPDSFNPMLTYLIDTTSPEIKLARSSQVFGVLYPAGATTSQDSVTDLFGKCL 120
```

Db 61 VLPPDSDFNPLMTLYLTDTTSPNEIKLARKSEVYVAVKLPAGATTNSODGVTDLLGKCL 120
Qy 121 PVLQEWENHMPLLVHGEVNTPEVDMFDERKVFIEITVLRPLVQKFPOLKVMHEHTTTIDA 180
Db 121 PVLQEWENHMPLLVHGEVNTPEVDMFDERKVFIEITVLRPLVQKFPOLKVMHEHTTTIDA 180
Qy 181 VKFVESCCEGVAATVTPQHLVLRNNSLFOGGLQPHNYCLPVLRKREIHRALVSATVSGS 240
Db 181 VKFVESCCEGVAATVTPQHLVLRNNSLFOGGLQPHNYCLPVLRKREIHRALVSATVSGS 240
Qy 241 KRFELGTDSAPHDRRREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFNGPDF 300
Db 241 KRFELGTDSAPHDRRREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFNGPDF 300
Qy 301 YGLPRNNSKIKLSTKTPKVPESFSYASGDIIPMFAGEMLDWLPAP 344
Db 301 YGLPRNNSKIKLSTKTPKVPESFSYASGDIIPMFAGEMLDWLPAP 344

RESULT 2

US-10-437-963-189313
/ Sequence 189313, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Bardazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ NUMBER OF SEQ ID NOS: 2003-05-14
/ SEQ ID NO 189313
/ LENGTH: 359
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_85834C.1.pap
US-10-437-963-189313

Query Match 80.5%; Score 1458; DB 16; Length 359;
Best Local Similarity 78.1%; Pred. No. 8.3e-140;
Matches 268; Conservative 39; Mismatches 36; Indels 0; Gaps 0;

Qy 2 ELSTTODDWHHLRBDGDLKAVVSHSAHFGRAIVMPLKPPITTTAAVAAYREAILKS 61
Db 12 ELSTTODDWHHLRBDGDLKAVVSHSAHFGRAIVMPLKPPITTTAAVAAYREAILKS 61
Qy 62 LPPVDSDFNPLMTLYLTDTTSPNEIKLARKSEVYVAVKLPAGATTNSODGVTDLLGKCL 120
Db 72 LPPVDSDFNPLMTLYLTDTTSPNEIKLARKSEVYVAVKLPAGATTNSODGVTDLLGKCL 120
Qy 122 VLQEWENHMPLLVHGEVNTPEVDMFDERKVFIEITVLRPLVQKFPOLKVMHEHTTTIDA 180
Db 122 VLQEWENHMPLLVHGEVNTPEVDMFDERKVFIEITVLRPLVQKFPOLKVMHEHTTTIDA 180
Qy 132 VLEEMARQEMPLLHGEVNTPEVDMFDERKVFIEITVLRPLVQKFPOLKVMHEHTTTIDA 191
Db 132 VLEEMARQEMPLLHGEVNTPEVDMFDERKVFIEITVLRPLVQKFPOLKVMHEHTTTIDA 191
Qy 182 KRFVESCCEGVAATVTPQHLVLRNNSLFOGGLQPHNYCLPVLRKREIHRALVSATVSGS 241
Db 182 KRFVESCCEGVAATVTPQHLVLRNNSLFOGGLQPHNYCLPVLRKREIHRALVSATVSGS 241
Qy 242 KRFELGTDSAPHDRRREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFNGPDF 301
Db 242 KRFELGTDSAPHDRRREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFNGPDF 301
Qy 301 YGLPRNNSKIKLSTKTPKVPESFSYASGDIIPMFAGEMLDWLPAP 344
Db 301 YGLPRNNSKIKLSTKTPKVPESFSYASGDIIPMFAGEMLDWLPAP 344

RESULT 3

US-10-424-599-149491
/ Sequence 149491, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ NUMBER OF SEQ ID NOS: 2003-04-28
/ SEQ ID NO 149491
/ LENGTH: 379
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(379)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_106012C.1.pap
US-10-424-599-149491

Query Match 80.2%; Score 1454; DB 15; Length 379;
Best Local Similarity 78.0%; Pred. No. 2.3e-139;
Matches 269; Conservative 36; Mismatches 40; Indels 0; Gaps 0;

Qy 1 MELSTTODDWHHLRBDGDLKAVVSHSAHFGRAIVMPLKPPITTTAAVAAYREAILKS 60
Db 34 MELSTTODDWHHLRBDGDLKAVVSHSAHFGRAIVMPLKPPITTTAAVAAYREAILKS 60
Qy 61 LPPVDSDFNPLMTLYLTDTTSPNEIKLARKSEVYVAVKLPAGATTNSODGVTDLLGKCL 120
Db 94 LPPVDSDFNPLMTLYLTDTTSPNEIKLARKSEVYVAVKLPAGATTNSODGVTDLLGKCL 120
Qy 121 PVLQEWENHMPLLVHGEVNTPEVDMFDERKVFIEITVLRPLVQKFPOLKVMHEHTTTIDA 180
Db 154 PVLQEWENHMPLLVHGEVNTPEVDMFDERKVFIEITVLRPLVQKFPOLKVMHEHTTTIDA 180
Qy 181 VKFVESCCEGVAATVTPQHLVLRNNSLFOGGLQPHNYCLPVLRKREIHRALVSATVSGS 240
Db 214 VKFVESCCEGVAATVTPQHLVLRNNSLFOGGLQPHNYCLPVLRKREIHRALVSATVSGS 240
Qy 241 KRFELGTDSAPHDRRREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFNGPDF 300
Db 274 KRFELGTDSAPHDRRREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFNGPDF 300
Qy 301 YGLPRNNSKIKLSTKTPKVPESFSYASGDIIPMFAGEMLDWLPAP 345
Db 334 YGLPRNNSKIKLSTKTPKVPESFSYASGDIIPMFAGEMLDWLPAP 345

RESULT 4

US-10-282-122A-48169
/ Sequence 48169, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zykkind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Yamamoto, Robert

```

APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 48169
LENGTH: 364
TYPE: PRT
ORGANISM: Burkholderia cepacia
US-10-282-122A-48169

```

```

Query Match 58.9%; Score 1067.5; DB 15; Length 364;
Best Local Similarity 60.4%; Pred. No. 6.3e-100;
Matches 206; Conservative 37; Mismatches 95; Indels 3; Gaps 2;

QY 3 LSTIQPDDMHLLRDGDVLAHVSHSAHFGRAIWMNLPKPTTTTAAVAAYEALIKSL 62
DB 22 LSLARPDDMHLLRDGDVLAHVSHSAHFGRAIWMNLPKPTTTTAAVAAYEALIKSL 81
QY 63 PVSDPFPMLTYLTDTTSPMEIKLARESGVGVKLYPAGATTNSDGVTDLFGKCLPV 122
DB 82 PAGMTFEPMLTYLTDTTSPMEIKLARESGVGVKLYPAGATTNSDGVTDLFGKCLPV 140
QY 123 LOEWEHNMPLLVHGEVTNPEVDMFDEKVFIEVLRPLVQKFPOLKVMEHVTITDAVK 182
DB 141 LEAQOETGMPLLVHGEVTNPEVDMFDEKVFIEVLRPLVQKFPOLKVMEHVTITDAVK 200
QY 183 FVE--SCTEGFVAATVTPQHLVLRNSLFOGGLQPHNYCLPVLKREIHRALVSAVTSGS 240
DB 201 YVRDADAAPGLIGCTITTAHLLVLRNMFVGGIRPHYCLPVLRKREHRAVALVEAASGN 260
QY 241 KRPFGLTDSAPHDRRRKESCGCAGIYNAPVALSVKAVEKENALDKLEAFISFNGPDF 300
DB 261 PRFPLGTDSAPHADAKETACGAGCYTALHLELYAEAFDTAGALDKLEGFASFQADF 320
QY 301 YGLPRNNSKITLSTKTPMKVPESEFSYASGDIIPMAGEMLDW 341
DB 321 YGLPRSAETVTLRREPWELPREIFAGETFPVPLRGEGITGW 361

```

```

APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 50426
LENGTH: 352
TYPE: PRT
ORGANISM: Burkholderia mallei
US-10-282-122A-50426

Query Match 58.4%; Score 1058; DB 15; Length 352;
Best Local Similarity 58.4%; Pred. No. 5.6e-99;
Matches 199; Conservative 45; Mismatches 95; Indels 2; Gaps 1;

QY 3 LSTIQPDDMHLLRDGDVLAHVSHSAHFGRAIWMNLPKPTTTTAAVAAYEALIKSL 62
DB 9 LTLARPDDMHLLRDGDVLAHVSHSAHFGRAIWMNLPKPTTTTAAVAAYEALIKSL 68
QY 63 PVSDPFPMLTYLTDTTSPMEIKLARESGVGVKLYPAGATTNSDGVTDLFGKCLPV 122
DB 69 PAGMTFEPMLTYLTDTTSPMEIKLARESGVGVKLYPAGATTNSDGVTDLFGKCLPV 128
QY 123 LOEWEHNMPLLVHGEVTNPEVDMFDEKVFIEVLRPLVQKFPOLKVMEHVTITDAVK 182
DB 129 LEAQOETGMPLLVHGEVTNPEVDMFDEKVFIEVLRPLVQKFPOLKVMEHVTITDAVK 188
QY 183 FVE--SCTEGFVAATVTPQHLVLRNSLFOGGLQPHNYCLPVLKREIHRALVSAVTSGS 240
DB 189 YVRDADAAPGLIGCTITTAHLLVLRNMFVGGIRPHYCLPVLRKREHRAVALVEAASGN 248
QY 241 KRPFGLTDSAPHDRRRKESCGCAGIYNAPVALSVKAVEKENALDKLEAFISFNGPDF 300
DB 249 PRFPLGTDSAPHADAKETACGAGCYTALHLELYAEAFDTAGALDKLEGFASFQADF 308
QY 301 YGLPRNNSKITLSTKTPMKVPESEFSYASGDIIPMAGEMLDW 341
DB 309 YGLPRSAETVTLRREPWELPREIDAGAGFPVPLRGEGITGW 349

```

```

RESULT 6
US-10-282-122A-77674
; Sequence 77674, Application US/10282122A

```

```

RESULT 5
US-10-282-122A-50426
; Sequence 50426, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haelebeck, Robert

```

```

Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EPIRA.034A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-09-06
PRIOR FILING DATE: 2000-09-06
PRIOR FILING DATE: 2000-11-27
PRIOR FILING DATE: 2000-11-27
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2001-02-09
PRIOR FILING DATE: 2001-02-09
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 77674
LENGTH: 350
TYPE: PRT
ORGANISM: Vibrio cholerae
US-10-282-122A-77674

Query Match
Best Local Similarity 57.4%; Score 1039.5; DB 15; Length 350;
Matches 200; Conservative 45; Mismatches 91; Indels 3; Gaps 3;
3 LSTTODDDHMLLRGQDYLKAVVSHSAHFGRAIWNPNKRPITTTAAVAAYRAIISL 62
12 LITTPDDHMLLRGQDYLKAVVSHSAHFGRAIWNPNKRPITTTAAVAAYRAIISL 62
63 PVDSPNDMLTYLTDTPSPMEIKLARESOVFGVKLYPAGATTNSODGVTDLFGCLPY 71
72 P-QAHFELMALYLTDTPSPMEIKLARESOVFGVKLYPAGATTNSODGVTDLFGCLPY 122
123 LOEVENHMLLVHGEVTPNEVDMFDRKVFIEYVLEPLVOKPOLKVMHEVHTTIDAVK 129
130 LOEVENHMLLVHGEVTPNEVDMFDRKVFIEYVLEPLVOKPOLKVMHEVHTTIDAVK 182
130 LOEVENHMLLVHGEVTPNEVDMFDRKVFIEYVLEPLVOKPOLKVMHEVHTTIDAVK 182
183 FVSCGEGVAATVTPQHLVLRNSLFOGLOPHNYCLPYLKEIHEREALVAALVSSGSKR 242
190 FVQOAGDN-VAAATTAHNLFFNRNMLVGGIRHFFCLPLKRAVCHQALVVAATISSSK 248
243 FFLGTDSPHRRRKEGSCGAGIYNAFVALSYAVFPEKENALDKLEAFISFNGEPFYG 302
249 FFLGTDSPHRRRKEGSCGAGIYNAFVALSYAVFPEKENALDKLEAFISFNGEPFYG 302
303 LPRNNSKIKLSKTWKVPSFVSAGDIIIMFAGEMLDW 341

```

```

Db
309 LPRNDETTLTKQAMPVASEMPSGSDIVPIRAGENIEM 347

RESULT 7
US-10-282-122A-51386
Sequence 51386, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EPIRA.034A
CURRENT FILING DATE: 2003-02-20
PRIOR FILING DATE: 2003-02-20
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-09-06
PRIOR FILING DATE: 2000-09-06
PRIOR FILING DATE: 2000-11-27
PRIOR FILING DATE: 2000-11-27
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2001-02-09
PRIOR FILING DATE: 2001-02-09
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 51386
LENGTH: 354
TYPE: PRT
ORGANISM: Bordetella pertussis
US-10-282-122A-51386

Query Match
Best Local Similarity 55.7%; Score 1008.5; DB 15; Length 354;
Matches 196; Conservative 46; Mismatches 97; Indels 5; Gaps 3;
2 ELSTTODDDHMLLRGQDYLKAVVSHSAHFGRAIWNPNKRPITTTAAVAAYRAIISL 58
8 ELSTTODDDHMLLRGQDYLKAVVSHSAHFGRAIWNPNKRPITTTAAVAAYRAIISL 58
59 LKSLPVD-SPDNDMLTYLTDTPSPMEIKLARESOVFGVKLYPAGATTNSODGVTDLFG 117
68 LRAAGDGTAAFTPLMTLYLTDTPSPMEIKLARESOVFGVKLYPAGATTNSODGVTDLFG 127
118 KCLPVLQEVNHNMLLVHGEVTPNEVDMFDRKVFIEYVLEPLVOKPOLKVMHEVHTT 127
128 KCGAALALEGCGMPLVHGEVTPNEVDMFDRKVFIEYVLEPLVOKPOLKVMHEVHTT 177
178 IDAVYFVSCGEGVAATVTPQHLVLRNSLFOGLOPHNYCLPYLKEIHEREALVAALV 237
188 REGAHVARD-AEGPTAATITPQHLVLRNSLFOGLOPHNYCLPYLKEIHEREALVAALV 246

```

```
QY 238 SGGKRFLLGTDSAPHDRRREKSCCGAGIYNAPVALSVYAKVPEKENALDKLEAFTSFG 297
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 247 SGGPRFLLGTDSAPHARGLKEHACGAGCTTAAHAMELVATAFDAGRLDRLEGFASFG 306
QY 298 PDYVGLPRNNSKIKLSTPKWKPESFSYASGDIIPMFAGEMLDW 341
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 307 PDYVGLPRNTGTITLRREAYEIPAEVAFGDTTIVPSGSESLW 350

RESULT 8
US-10-282-122A-65811
/ Sequence 65811, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 65811
/ LENGTH: 344
/ TYPE: PRT
/ ORGANISM: Neisseria meningitidis
US-10-282-122A-65811

Query Match 54.2%; Score 983; DB 15; Length 344;
Best Local Similarity 56.3%; Pred. No. 2.4e-91;
Matches 192; Conservative 55; Mismatches 88; Indels 6; Gaps 5;

QY 3 LSTQPDWMHLRLRGDYLKAVVSHAHFGRALVMPNLKPPITTTAAVAVREAILKSL 62
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 LTIIRPDMHLRLRGDYLKAVVSHAHFGRALVMPNLKPPITTTAAVAVREAILKSL 63
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 PVDSDFNPLMTLYLTDTTSPMEIKLARSSQVVGKLYPAGATTNSQDGVTDLFQKCLPV 122
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 64 PEGSAFEPMLTLYLTDOATPALVREAKAAIV-AFKLYPAGATTNSDGVTDLF-KLIPV 121
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 123 LOENVEHNMPLLVHSEVTNPEVDMDREKVFIETVLRPLVQCFPOLKVMMEVITTIIDAVK 182
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 122 LBEAMAKOGLFLVHGEVTDPEIDIPDREAPFIGRWKPVLAQVPMNLKVFEHITTAENAR 181
QY 183 PUESCEGFAATVTPQHVLINRNSLFOGDLDPHNYCLPVLRREIRALVSATSGSK- 241
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 182 LVLEAGDN-VAATVTPQHLLNRNDLVGVRPHHCLPVLRREIRALVSATSGSK- 239
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 -PFFLTGDSAPHDRRREKSCCGAGIYNAPVALSVYAKVPEKENALDKLEAFTSFGPDF 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 HKFFLTGDSAPHARGLKEHACGAGMFSMTAIELVAEVPKXGALDKLEAFASKNGARF 299
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 301 YGLPRNNSKIKLSTPKWKPESFSYASGDIIPMFAGEMLDW 341
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 YGIPENTDTITLVKQQTVPASIPYGDGLVPMRAGGEIGW 340

RESULT 9
US-10-282-122A-65238
/ Sequence 65238, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 65238
/ LENGTH: 344
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-282-122A-65238

Query Match 54.0%; Score 979; DB 15; Length 344;
Best Local Similarity 56.3%; Pred. No. 6.2e-91;
Matches 192; Conservative 55; Mismatches 88; Indels 6; Gaps 5;

QY 3 LSTQPDWMHLRLRGDYLKAVVSHAHFGRALVMPNLKPPITTTAAVAVREAILKSL 62
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 LTIIRPDMHLRLRGDYLKAVVSHAHFGRALVMPNLKPPITTTAAVAVREAILKSL 63
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
QY 63 PVDSDPNPLMTLYLTDTTSPMEIKLARESOVFGVKLYPAGATTNSODGVTDLFGKCLPV 122
Db 64 PGCSAFBPLMTLYLTDOQTPBELVRAKAAIGIV-APKLYPAGATTNSGSGVTDLF-KLIPV 121
QY 123 LOEMENHMPLLVHGEVTNPEVDMFDRKVFIEYTLRPLVQKFPOLKVMEHTTTIDAVK 182
Db 122 LERMEKIGMPLVHGEVTNPEVDMFDRKVFIEYTLRPLVQKFPOLKVMEHTTTIDAVK 182
QY 183 FVESCIEGFVAATVPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSGSKR 241
Db 182 LVLEAGDN-VAATVPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSGSKR 239
QY 242 -RFLGTDSPHRRRREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSPGDPFG 300
Db 240 HKFGLTDSAPHARREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSPGDPFG 300
QY 301 YGLPRNNSKIKLSTPKVPEFSFYASGDIIPMFAGEMLDW 341
Db 300 YGIPENADTITIVQSGQTVASVPYGDGELVPRAGGETGW 340
```

RESULT 10

```
US-10-282-122A-59728
/ Sequence 59728, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Travick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 59728
/ LENGTH: 348
/ TYPE: PRT
/ ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59728
```

```
Query Match          54.0%; Score 978; DB 15; Length 348;
Best Local Similarity 55.8%; Pred. No. 8e-91;
```

```
Matches 189; Conservative 46; Mismatches 102; Indels 2; Gaps 2;
QY 3 LSTIQPDWHLHRLDGDVLKAVVSHSAHFGRAIWNPNKPRITTTAAVAAREAILKSL 62
Db 8 LKIRRPDMHILRRDDMLKTVTPYTSFYGALVWNPVPTVVAIAAIAKORIMAV 67
QY 63 PVDSDPNPLMTLYLTDTTSPMEIKLARESOVFGVKLYPAGATTNSODGVTDLFGKCLPV 122
Db 68 PACHDFPLMTLYLTDTTSPMEIKLARESOVFGVKLYPAGATTNSODGVTDLFGKCLPV 126
QY 123 LOEMENHMPLLVHGEVTNPEVDMFDRKVFIEYTLRPLVQKFPOLKVMEHTTTIDAVK 182
Db 127 LERMEKIGMPLVHGEVTNPEVDMFDRKVFIEYTLRPLVQKFPOLKVMEHTTTIDAVK 186
QY 183 FVESCIEGFVAATVPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSGSKR 242
Db 187 YVADGNE-LLAATITPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSGSKR 245
QY 243 FFLGTDSPHRRRREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSPGDPFG 302
Db 246 AFLGTDSPHRRRREKSCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSPGDPFG 305
QY 303 LPRNNSKIKLSTPKVPEFSFYASGDIIPMFAGEMLDW 341
Db 306 LPRNNSKIKLSTPKVPEFSFYASGDIIPMFAGEMLDW 344
```

RESULT 11

```
US-10-282-122A-66489
/ Sequence 66489, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Travick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 66489
/ LENGTH: 348
```



```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66489

Query Match      53.8%; Score 975; DB 15; Length 348;
Best Local Similarity 55.5%; Pred. No. 1.6e-90;
Matches 188; Conservative 43; Mismatches 106; Indels 2; Gaps 2;

QY 3 LSTGDDMHHLRDGDLKAVVSHSNAHFGRAIWMENLKPITTTAAVAAYEAILKSL 62
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 5 LTLRRDDMHHLRDGAALANTVGDAARTGRAIWMENLVPYRNAEADAYQRIILAAQ 64
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 63 PVDSDFNPLMTLYLTDTTSPMEIKLARESQVGVKLYPAGATTNSGQVTDLFGKCLPV 122
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 65 PAASREPLWVLYLTTRTSTETRIKAKSGFVAAKLYPAGATTNSGQVTRI-DNIFEA 123
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 123 LQEMVHNHNPVLVHGEVTNPEVDMFDEKVFIEFTVLRPLVQKFPOLKVMEHTTTIDAVK 182
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 124 LEMAAEGMPLVHGEVTRAEVDVDFREKQFIDENLRVRVERPPTLKVFEHITTDGAAQ 183
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 183 FVESCIEGFVAATVTPQHLVLRNNSLFQGGLOPHNYCLPVLKREIHREALVSAVTSGSKR 242
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 184 FVNEAPRN-VGATTTAHLVYNNHMLVGGIRPHFYCLPLTKGTHQEDALLDAVSGNPK 242
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 243 FFLGTDSAPHRDRRKESCCGAGIYNAPVALSVYAKVFEKENALDKLEATFSFGPDPFG 302
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 243 FFLGTDSAPHRARHKEAACGACGYSAYAAIELYAEAFEOQNALDKLEGFASLHGPDPFG 302
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 303 LPRNNSKIKLSKTPWKVPSFSYASGDIIPFAGEMLDW 341
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 303 LPRNTDRITLVREEMQAPASLPFGDPVDFLRAGETLRW 341
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 12
US-10-282-122A-43220
; Sequence 43220, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
```

```
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43220
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43220

Query Match      53.6%; Score 971; DB 15; Length 348;
Best Local Similarity 55.8%; Pred. No. 4.2e-90;
Matches 189; Conservative 46; Mismatches 102; Indels 2; Gaps 2;

QY 3 LSTGDDMHHLRDDVDLKAVVSHSNAHFGRAIWMENLKPITTTAAVAAYEAILKSL 62
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 8 LKLRDDMHHLRDDMLTVPYSEIYGRAIWMENLAPVTVYEAAYHQRILLDAV 67
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 63 PVDSDFNPLMTLYLTDTTSPMEIKLARESQVGVKLYPAGATTNSGQVTDLFGKCLPV 122
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 68 PACHDTPMLTCYLTSLDNELERGFNEGVFTNAKLYPANATTNSHGVTSTI-DAIMPV 126
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 123 LQEMVHNHNPVLVHGEVTNPEVDMFDEKVFIEFTVLRPLVQKFPOLKVMEHTTTIDAVK 182
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 127 LEMEKIEMPLVHGEVTHADIDIFDEARFISVMEPLRQRLTALKVFEHITTKDAAD 186
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 183 FVESCIEGFVAATVTPQHLVLRNNSLFQGGLOPHNYCLPVLKREIHREALVSAVTSGSKR 242
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 187 YVADGNER-LAAITTPQHLMFNRRHMLVGGVRPHLYCLPLTKNTHOQALRELVASGPNR 245
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 243 FFLGTDSAPHRDRRKESCCGAGIYNAPVALSVYAKVFEKENALDKLEATFSFGPDPFG 302
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 246 VFLGTDSAPHRARHKESSCCGACGFNAPVALSVYAKVFEKENALDKLEATFSFGPDPFG 305
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 303 LPRNNSKIKLSKTPWKVPSFSYASGDIIPFAGEMLDW 341
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 306 LPRVNDTFIELVREEQVAESIALTDDTLVFLAGETVRW 344
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 13
US-10-282-122A-67595
; Sequence 67595, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
```

```

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67595
; LENGTH: 348
; TYPE: PRF
; ORGANISM: Pseudomonas putida
US-10-282-122A-67595
```

```

Query Match      53.1%; Score 962; DB 15; Length 348;
Best Local Similarity 53.7%; Pred. No. 3,5e-89;
Matches 182; Conservative 46; Mismatches 109; Indels 2; Gaps 2;
```

```

QY 3 LSIITPPDDWHLRLRDGDVLKAVSHSAHFGRAIVMPNLKPPITTTAAVAAYREALIKSL 62
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 5 LTLRLPDDWHLRLRDGAVLPHTVGDVARTFARAILMPVLVPVNAQADVIRGRIILAA 64
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 63 PVDSDPNPLMTLYLNDTTSPEMIKARSOVFGVKLYPAGATTNSODGVTDLFCKCLPV 122
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 65 PAGESFPELMVLYLTDRTSPEDVRAAKASGFVYAAKLYPAGATTNSDGVTSI-DKIFPA 123
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 123 LQEWENHNPPLLVHGEVTPNPEVDMFDEKVFIEFTVLRPLVOKPOLKVMHEHTTIDAVK 182
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 124 IEALAEVGMPLLVHGEVTRSEIDVDRKRIIDENMRVVERFPTLVVEHITTSDAQ 183
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 183 FVESTGEFVAATYTPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSQSKR 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 184 FVTEAPAN-VGATITTAHLVLRNRMVLVGGIRPFFCYLPILKRWTHQVALLDAATSGNPK 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 243 FFLGDSAPHDRRRCSCGCGAGIYNAPVALSVYAKFEKENALDKLEATTSNGDPFYG 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 243 FFLGDSAPHDRRRCSCGCGAGIYNAPVALSVYAKFEKENALDKLEATTSNGDPFYG 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 303 LPRNNSKIKLSKTPWKVPESFVSAGDIIPMFAGEMLDW 341
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 303 LPRNNSKIKLSKTPWKVPESFVSAGDIIPMFAGEMLDW 341
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 303 LPRNNSKIKLSKTPWKVPESFVSAGDIIPMFAGEMLDW 341
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 14

```

US-10-282-122A-69927
; Sequence 69927, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
```

```

; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69927
; LENGTH: 347
; TYPE: PRF
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69927
```

```

Query Match      52.9%; Score 959; DB 15; Length 347;
Best Local Similarity 55.2%; Pred. No. 7e-89;
Matches 187; Conservative 44; Mismatches 106; Indels 2; Gaps 2;
```

```

QY 3 LSIITPPDDWHLRLRDGDVLKAVSHSAHFGRAIVMPNLKPPITTTAAVAAYREALIKSL 62
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 5 LTLRLPDDWHLRLRDGAVLPHTVGDVARTFARAILMPVLVPVNAQADVIRGRIILAA 64
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 63 PVDSDPNPLMTLYLNDTTSPEMIKARSOVFGVKLYPAGATTNSODGVTDLFCKCLPV 122
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 65 PAGESFPELMVLYLTDRTSPEDVRAAKASGFVYAAKLYPAGATTNSDGVTSI-DKIFPA 123
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 123 LQEWENHNPPLLVHGEVTPNPEVDMFDEKVFIEFTVLRPLVOKPOLKVMHEHTTIDAVK 182
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 124 IEALAEVGMPLLVHGEVTRSEIDVDRKRIIDENMRVVERFPTLVVEHITTSDAQ 183
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 183 FVESTGEFVAATYTPQHLVLRNSLFOGGLQPHNYCLPVLRKRIHREALVSAVTSQSKR 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 184 FVTEAPAN-VGATITTAHLVLRNRMVLVGGIRPFFCYLPILKRWTHQVALLDAATSGNPK 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 243 FFLGDSAPHDRRRCSCGCGAGIYNAPVALSVYAKFEKENALDKLEATTSNGDPFYG 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 243 FFLGDSAPHDRRRCSCGCGAGIYNAPVALSVYAKFEKENALDKLEATTSNGDPFYG 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 303 LPRNNSKIKLSKTPWKVPESFVSAGDIIPMFAGEMLDW 341
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 303 LPRNNSKIKLSKTPWKVPESFVSAGDIIPMFAGEMLDW 341
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 15

```

US-10-282-122A-78404
; Sequence 78404, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 06:47:55 ; Search time 29 Seconds
(without alignments)
890.641 Million cell updates/sec

Title: US-10-070-277-2

Perfect score: 1812
Sequence: 1 MELSTQDDHHLRDGDV.....SGDIIPMFAGEMLDWIPAPL 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgm2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgm2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgm2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgm2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgm2_6/prodata/1/iaa/PCITUS_COMB.pep:*
6: /cgm2_6/prodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	978	54.0	352	4	US-09-489-039A-7589 Sequence 7589, App
2	975	53.8	398	4	US-09-252-991A-17491 Sequence 17491, A
3	952.5	52.6	346	4	US-09-328-352-5541 Sequence 5541, App
4	946	52.2	355	4	US-09-540-236-5330 Sequence 2330, App
5	930	51.3	356	4	US-09-543-681A-4285 Sequence 4285, App
6	522.5	28.8	365	4	US-09-248-796A-18106 Sequence 18106, A
7	132	7.3	734	4	US-09-902-540-11399 Sequence 11399, A
8	119.5	6.6	430	4	US-09-107-532A-5107 Sequence 5107, App
9	116.5	6.4	448	4	US-09-134-000C-6087 Sequence 6087, App
10	103.5	5.7	422	4	US-09-583-110-3844 Sequence 3844, App
11	103.5	5.7	426	4	US-09-107-433-3349 Sequence 3349, App
12	102.5	5.7	371	4	US-09-710-279-628 Sequence 628, App
13	102.5	5.7	425	4	US-09-710-279-2604 Sequence 2604, App
14	102.5	5.7	427	3	US-09-134-001C-4803 Sequence 4803, App
15	97	5.4	244	3	US-08-469-318-126 Sequence 126, App
16	97	5.4	244	3	US-08-468-609A-126 Sequence 126, App
17	97	5.4	244	3	US-08-446-872A-126 Sequence 126, App
18	97	5.4	244	4	US-08-762-227A-126 Sequence 126, App
19	97	5.4	244	5	PCT-US95-01185-126 Sequence 126, App
20	96.5	5.3	699	4	US-09-438-185A-506 Sequence 506, App
21	94	5.2	244	3	US-08-469-318-125 Sequence 125, App
22	94	5.2	244	3	US-08-468-609A-125 Sequence 125, App
23	94	5.2	244	3	US-08-446-872A-125 Sequence 125, App
24	94	5.2	244	4	US-08-762-227A-125 Sequence 125, App
25	94	5.2	244	5	PCT-US95-01185-125 Sequence 125, App
26	94	5.2	285	3	US-08-875-533-69 Sequence 69, App
27	89.5	4.9	681	4	US-09-252-991A-18889 Sequence 18889, A

28	89.5	4.9	734	4	US-09-809-665A-117 Sequence 117, App
29	89	4.9	244	3	US-08-469-318-127 Sequence 127, App
30	89	4.9	244	3	US-08-468-609A-127 Sequence 127, App
31	89	4.9	244	3	US-08-446-872A-127 Sequence 127, App
32	89	4.9	244	4	US-08-762-227A-127 Sequence 127, App
33	89	4.9	244	5	PCT-US95-01185-127 Sequence 127, App
34	89	4.9	285	4	US-08-875-533-70 Sequence 70, App
35	89	4.9	458	4	US-09-497-585A-8 Sequence 8, App
36	89	4.9	458	4	US-09-916-501A-14 Sequence 14, App
37	89	4.9	469	4	US-09-540-236-2593 Sequence 2593, App
38	88.5	4.9	627	4	US-09-360-545-32 Sequence 32, App
39	88.5	4.9	868	4	US-09-949-016-11723 Sequence 11723, A
40	88	4.9	191	2	US-08-438-439C-22 Sequence 22, App
41	88	4.9	244	3	US-08-469-318-136 Sequence 136, App
42	88	4.9	244	3	US-08-468-609A-136 Sequence 136, App
43	88	4.9	244	3	US-08-446-872A-136 Sequence 136, App
44	88	4.9	244	4	US-08-762-227A-136 Sequence 136, App
45	88	4.9	244	5	PCT-US95-01185-136 Sequence 136, App

ALIGNMENTS

```
RESULT 1
US-09-489-039A-7589
Sequence 7589, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489, 039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117, 747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7589
LENGTH: 352
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7589

Query Match      54.0% Score 978; DB 4; Length 352;
Best Local Similarity 55.8% Pred. No. 1.8e-98;
Matches 189; Conservative 46; Mismatches 102; Indels 2; Gaps 2;

QY 3 LSTQDDHHLRDGDVLRKAVVSHSHHGRAIVNPNLKPITTTAAVAVERAIKSL 62
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 12 LKTRPDDHHLRDDMLKTVVPYSEFGRAIVNPNLVPPTTAAVAIRQIMDAV 71
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 PVSDPFLMTLYLTDTTSFMEIKLAESQVVFVKVLPAGATTNSQDVTDLFGKCLPV 122
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 72 PACHDFPLMTCTLTLDLPAELERGFNEGVFTAAKLYPANATNSHGVST-DAIMPV 130
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 123 LQEMVENHMLVHGEVTNEVDMFDEKPIETIVLRPLVOKFPQLVWNEHVTIDAVK 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 131 LERNEKGMPLVHGEVTHAEIDIFDEKAFIETVMEPLKORLGLVVEHITTKDAE 190
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 183 FVDSCTGEFAATVTPQHLVLRNSLFGGLQHPNNYCLPYLKEIRHREALVSATSGSKR 242
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 191 YKRDGNL-LIAATITPQHLVFNRRHMLVGRIHRLVCLPVLKXNIHQALRELVASGFSR 249
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 243 PFLQTSAPDRRRKESCCGACIYNAPVALSVYAKVFEKENALDKLEAFISFGPDPIYG 302
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 250 AFLQTSAPDRRRKESCCGACIYNAPVALSVYAKVFEKENALDKLEAFISFGPDPIYG 309
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 303 LPRNNSKIKSKTPKWKVPSFSYASGDIIMFAGEMLDW 341
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 310 LPVNESYVELVRETTVDVSDIALPNDTLVPLAGETVRW 348
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 2

```

Db      6  ITLLQPDMDWHALIRDGIALKRTVPBPLAQFARAIQMPVLVPEPVKTVEEALAYRREILIANH 65
QY      63  PYVDSDFNLMTLYLIDTTS.PMEIKLARESOVFGKLYBPATATNSODGTYDLPGKCLPY 12
Db      66  PEGNNFDPBMYLFTDHTSPDEAVRKIKSEHVMYIKLYPATATNSDGVSDI-RKYAV 12
QY      123 LOEMVENHMLPLVGEYTNPEVDMFDEKXYILETVAPLVOKFPOLKVMMEVTTIDAVK 18
Db      125 IEQLEEHQVPLPLHGEVTHNHVDLFDEKREKLDEVLSPBLKQFPLKLVYEHITSDAAH 18
QY      183 FVECTSECFVAATYTPQHLVLNRNSLFOGQIQPNYCLPYVLKREIIRREALVSAVTSQSK 24
Db      185 FVLV-QQBNVAATITPQHLFENRNDIMVGKIKPFIYCLPIIKRQTHQITLLEVAISGNPK 24
QY      243 FLIGTDSAPHDRREKESCGCAGLYNAPVLVSYAKVFEKNMLDKLEAFTSPGDPFYG 30
Db      244 FFLIGTDSAPHAQNAKENACGAGCYSPAPNILELYAQNFDQVGLERIEGFASHFGADFYG 30

```

Db 304 LPRNTSTITLVKEDNLVPSFDYLDNOKIIPHAGETLQW 343

```

RESULT 4
US-09-540-236-2330
/ Sequence 2330, Application US/09540236
/ Patent No. 6673910
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAMPA
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2005-001
/ CURRENT APPLICATION NUMBER: US/09/540.236
/ CURRENT FILING DATE: 2000-04-04
/ NUMBER OF SEQ ID NOS: 3840
/ SEQ ID NO 2330
/ LENGTH: 355
/ TYPE: PRT
/ ORGANISM: M.catarrhalis
US-09-540-236-2330

Query Match      52.2%   Score 946;   DB 4;   Length 355;
Best Local Similarity 53.4%   Pred. No. 5.9e-95;
Matches 187;   Conservative 50;   Mismatches 97;   Indels 16;   Gaps 6;

QY      3      LSTIOPDDWHILHLDGCVLAAVSHSAHFGRAIVMNLKPITTTAAVAAYEEALIKSI 62

```

```

OY 63 PVDSD-----FNPLMTLYLTDTTSSMELKAR:SESQVNVGVLYLPAGATTNQOGYTD 114
Db 66 K-DSDLPATRKDSFDPFRNVLYLTQTTAKDIDDAQIGIVSAVLYPPGATTNADGYTD 124
OY 115 LFGKCLPILQDMVBNHNPVLVHGEVTNPEVDMFDERKVFJETVLRPLVQKPEPOLKVMVH 174
Db 125 ILARSL-VFEMQGHGQIPILVHGEITSNIDIFDERKRPDDVLNKKIITQEPNLIKVMH 183
OY 175 VTTIDAVKVFESCTE--GFVAAVTPPHLYLVNRSLSFQGLQPHNYCYLPVKGEIHEAL 232
Db 184 ITTMDADP---CLAQGNHTAATITTPHQLLFNRRHLLLVGSIKHYCYLPILKRAAQKRL 240
OY 233 VSAVTSSGSKRFPJGTSAPHRDRRKESCCGAGIYNAPVLSYAKVFEEKMLDKEAF 292
Db 241 LEVATSGNPKREFJGTGDPAPHTHTKSSCCGACACISALHLPYNAAFESVNALDKEFN 300

```

RESULT 5

AS-09-543-681A-4285

Dd 301 ASRYGAQFGLAINTQTILIKQPITPSEFHYLGKNTLPLLAGVLPW 350

wx 293 ISFGGDFRIGLPPNNNSKITKLSTPMKWPEBFSY-ASGDIIIMPAGEMLDW 341

```
/ Sequence 4285, Application US/09543681A
/ Patent No. 6605709
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT APPLICATION NUMBER: US/09/543,681A
/ CURRENT FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 4285
/ LENGTH: 356
/ TYPE: PRT
/ ORGANISM: Proteus mirabilis
US-09-543-681A-4285

Query Match      51.3%; Score 930; DB 4; Length 356;
Best Local Similarity 52.8%; Pred. No. 3,4e-93;
Matches 179; Conservative 51; Mismatches 107; Indels 2; Gaps 2;

QY      3  LSTGDDMHLLHRRDGDVLAHVSHSAHFGRAIWMPLKRPITTTAAVAAYREALIKSL 62
      15  LTRRPDDMHVHRRDDMLKTVPYTSRYGRALVWPNLVPTTTEAARRYDRIKAAI 74
      63  PVSDENPLMTLYLDTTSPMEIKLARESQVFGVKLYPAGATTNSQDVTDLFGKCLPV 122
      75  PSGDKFEPLMTGCLTSTSEVGEQGLQGVFTACKLYPANATNSHGVSID-NKIYPI 133
QY      123  LQEMVHNHMLVHGEVTNEVDMFDEKRFIEVLRPLVQKRPOLKVMNEHTTIDAVK 182
      134  LSYMEKIGMPLIHGEVTASDIDIFDEARFIDNVAPVKPALKIVEHITTKEMAQ 193
QY      183  FVSSCTEGFAATVPOHLVLNNSLQGGLOPHNYCLPVLKCEIHREALVSAVTSQSKR 242
      194  YVLEGNE-FLGATITTPQHLNFNRHMLVGVKPHLYCLPLKRNVOEALROVAASHSR 252
QY      243  FLIGTDSAPHRRRKCCSCGAGIYNAPVALSVYAKVEKENALDKLEAFSPNGDPYFG 302
      253  FLIGTDSAPHLQRRKSSCCGAGVFNAPFTLAAVASFKELNLSHEAFCSLNGPRFYG 312
QY      303  LPRNNSKIKLSTKTPWKVPESFSYASGDIIIPFAGEMLDW 341
      313  LPVNEGTTITLTKSVTAPAEIMSGDEALIFLANEDITHW 351
DB

RESULT 6
US-09-248-796A-18106
/ Sequence 18106, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 18106
/ LENGTH: 365
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-18106

Query Match      28.8%; Score 522.5; DB 4; Length 365;
Best Local Similarity 38.3%; Pred. No. 1.6e-48;
Matches 139; Conservative 57; Mismatches 132; Indels 35; Gaps 15;
```

```
QY      1  MELSTIQDDMHLLHRRDGDVLAHVSHSAHFGRAI--VMPNLKRPITTTAAVAAYREALI 58
      11  IELGITA--DMVHLRDSMCE-LITPTRTGGIALSYVMPNLVPPITTKQGVVNHKTL 67
QY      59  LKSLPVSDENPLMTLYLDTTSPMEIKLARE-SQVFGVKLYPAGATTNSQDGV-TDLF 116
      68  SKLAP-QTTF--LMSFYLKDLTP---ELIEECADLIHGVCYCPAGVTTNSKGVDPNDF 121
QY      117  GKCLPVQEMVHNHMLVHGE---VMPVDMFDEKRFIEVLRPLVQKRPOLKVM 172
      122  SSFYPLEFMQKGYVLNIHGEKPTGSEDEINVAEPFL-PALKULHADEPRKLTIL 180
QY      173  EHVTITDAVKFVESCT-----EGFAATVTPQHLVLNNSLQGGLOPHNYCLPVLR 225
      181  EHCTSDAVEYTRBELNSNKPQGDLELVGATITTAHHYLIIDN-WAG--NPINFCKVAKF 237
QY      226  EIHREALVSAVTSQSKRFFLGTDSPADPHRRKCCSCG-CAGIYNAPVALSVAKVEKEN 284
      238  QDKRKALVDAATSGEPMFFGSDADPHIAKQTHVGACAGVYQSHALAYVAEIDPKAN 297
QY      285  ALDKLEAFSPNGDPYGL-----PRNNSKIKLSTKTPWKVPESFSYASGDIIIPFAGEM 338
      298  KLDNLKKFVSDNGIKFYGLSDILKKHKNESVWLVRDNQVPEIIGNGDVEVPFFRAGET 357
QY      339  LDW 341
      358  LKY 360
DB

RESULT 7
US-09-902-540-11399
/ Sequence 11399, Application US/09902540
/ Patent No. 6833447
/ GENERAL INFORMATION:
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Wiegand, Roger C.
/ TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
/ FILE REFERENCE: 38-10(15849)B
/ CURRENT APPLICATION NUMBER: US/09/902,540
/ CURRENT FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: 60/217,883
/ PRIOR FILING DATE: 2000-07-10
/ NUMBER OF SEQ ID NOS: 16825
/ SEQ ID NO 11399
/ LENGTH: 734
/ TYPE: PRT
/ ORGANISM: Myxococcus xanthus
US-09-902-540-11399

Query Match      7.3%; Score 132; DB 4; Length 734;
Best Local Similarity 22.5%; Pred. No. 2.9e-05;
Matches 81; Conservative 46; Mismatches 115; Indels 118; Gaps 18;

QY      10  DMHHLRD-GD-----VLKAVSHSAHFGRAIWMPLKRPITTTAAVAAYREALIKSLP 63
      359  DLHHLRPEEGEKETVLTGCRAAVAGGTAIVAMPYTK----- 397
QY      64  VDSDFNPLMTLYLDTTSPMEIKLARESQVFGVKLYPAGATTNSQDGV----- 111
      398  -----VVNDNAMTVELVLR-ARAAGLCHVYPAGAITGLGABELAEWGEIVSA 445
QY      112  ---VTLFGKCLPV-----LQEMVHNHMLVH----- 136
      446  GCAVITD-DGR--FVNNASLMRRTLQYATQFDVPMVHEEDLTSAGANHEGTTSTRLG 502
QY      137  --GEVTPPEVDMFDEKRFIEVLRPLVQKRPOLKVMNEHTTIDAVKVV-ESCTEGF-V 192
      503  LRGIAPAEVAMVARDVLEET-----KGRLVAAVSCGSVRLINEAKRSGLRV 553
QY      193  AATVTPQHLVLNNSLQGGLOPHNYCLPVLKREIHREALVSAVTSQSKRFFLGTDSPVH 252
```

Db 554 TCEVAPHHFIIDRAV--GDYDTHAKMAPLADTDVQALREALVDGTV--A1ATDAPH 610
QY 253 D--RRREKSCGAGIYNAPVALS-----VYAKVEKENALDKLEAFTSPGPR-DYGLP 304
Db 611 GVDKQVEKFGINGIVGLETALGLTELHVAGVLTARAVELLT-----HGPAKAGLP 665

RESULT 8

US-09-107-532A-5107
; Sequence 5107, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; APPLICATION DATA:
; APPLICATION NUMBER: 60/095,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...430
; SEQUENCE DESCRIPTION: SEQ ID NO: 5107:
US-09-107-532A-5107

Query Match
Best Local Similarity 23.3%; Score 119.5; DB 4; Length 430;
Matches 67; Conservative 43; Mismatches 114; Indels 63; Gaps 14;

QY 10 DMHLTRD-GDYLKAVV-----SHSAHFGRATVMPNKP-PITTTAAVAY-----REAI 58
Db 62 DVVHHRFBPGFTYKETIEAGTSARGGFTTVCAMPNINPVDTAETKLROYTDIIRKDAV 121
QY 59 LKSLPVDSDFNPLMTLYLDTTSPMEIKLARSQVVGKLYPAGATTNSQDGV-TDLFG 117
Db 122 VKVLQY-----APTENTL-RSEKLVQDEALIERGAPAFINDGVQYTAG 164
QY 118 KCLPVLQEVNHNMLPLVHGEVTNPEVDMF-----DREKVFIEVLRP-----160
Db 165 TMYLAKKEAANKNKLVANH--TEDESLRFGCVVHAKKAELGLPGILSTVESSQIARD 221

QY 161 -LVOKFPQKVMENHTTIDAYKVESCTEG--FVAATVTPQHLVLRNSLFGQGLQPHN 217
Db 222 LLLMAGCVHNVHCVSTKESVRIADKAKGIVHVAEVSPPHLLIDEDI-----PBD 275
QY 218 YCL-----PVLKREIHRREALVSAVTSKRFPLGTDSPAPDRRKXCS 260
Db 276 FGFWKMPPLRGREDEALIEGLDGTID-CIATDHAHPGLEKXQS 321

RESULT 9

US-09-134-000C-6087
; Sequence 6087, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 6087
; LENGTH: 448
; TYPE: PR
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6087

Query Match
Best Local Similarity 6.4%; Score 116.5; DB 4; Length 448;
Matches 80; Conservative 39; Mismatches 124; Indels 87; Gaps 18;

QY 10 DMHLTRD-GDYLKAVV-----SHSAHFGRATV-MPNKP-PITTTAAVAY-----REAI 58
Db 79 DVVHHRFBPGFTYKETIEAGTSARGGFTTVCAMPNINPVDTAETKLSEVYDLQKDAV 138
QY 59 LKSLPVDSDFNPLMTLYLDTTSPMEIKLARSQVVGKLYPAGATTNSQDGV-TDLFG 117
Db 139 VKVLQYAPITEELRSEVLINOKA-----LKEAR-----AFAFN--DGVGVQYTAG 181
QY 118 KCLPVLQEVNHNMLPLVHGEVTNPEVDMFDEKVFJETVLRPLVOK---FPOLKVM-- 173
Db 182 TMYLAKKEAANKNKLVANH--TEDESLRFGCVVHAKKAELGLPGILSTVESSQIARD 221
QY 174 -----HTTIDAVKVESCTEG--FVAATVTPQHLV-LRNSLFG 210
Db 233 SQIARDITLAKETGVHNVHCVSTKESVRIADKAKGIVHVAEVSPPHLLIDEDI 292
QY 211 GGLQPHNVCPLVLRKREIHRREALVSAVTSKRFPLGTDSPAPDRRKXCSGCGAGIYNAP 270
Db 293 BEFWKMN--PPLGLADRLQALIDGLDGTID-CIATDHAHPGLEKXQS-----PLANP 343
QY 271 VAL-----SVYAKVEKENALD 287
Db 344 FGIVGSETAFQLIYTNFVETGIFTLBQVID 373

RESULT 10

US-09-583-110-3844
; Sequence 3844, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATNO-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30

PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3844
LENGTH: 422
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3844

Query Match 5.7%; Score 103.5; DB 4; Length 422;
Best Local Similarity 20.8%; Pred. No. 0.017;
Matches 71; Conservative 54; Mismatches 131; Indels 85; Gaps 16;

10 DMHLHRLD-----GDVLKAVSHSAHFGRAIVMPNLKPPITTTAAVAAYREALIKS-- 61
57 DIHVHREBPQTHKEDIHTGALAAAGFTTVMMANTSPITSDVETLQAVLSAKKEI 116
62 -----LPVDSDFNPLMTLYITDFTTSPMEIKLARESOVFGVKLYPAGATTNSODGVTDLF 116
117 NVKTVATITGNFN---GKNLTDFKALLE-----AGAVGFSDDGIPLES 156
DB
117 GKCLPVLQGVENVHNMPLVHGEVTNPVDMFDEK-----VFIEVLRP----- 160
DB 157 SK---LVKEMAEAKKL-----NTFISLHEBDPGLNGVLGFENINAREHFIHICGATGV 206
QY 161 -----LVQKFPOLKVMHEHTTIDAVKFVESCTEGF---VAATVTPHVLNLRNS 207
DB 207 AEYAMAMARDMIAVATGAHIIQILSKESVRYE-FAQGLGAETVAEVAPOHSTKEAL 265
QY 208 LRF-QGGLQPHNYCLPVLKREIHRBALVSATSGSKREFLTGDSAPH--DRRKECGCGC 263
DB 266 LTRQGSNAKKN---PRLRLSDRRAVIEGLKSGVIT-VIATDHAIPHVDKVEDITKAP 321
QY 264 AGIYMAFVALSV--YAKVFEKENALDKLEAFISFNGPDY 301
DB 322 SGMTGLTSLSLGLTYLVEAGELSLMELBKMT-YNPAKLY 361

RESULT 11
US-09-107-433-3349
Sequence 3349, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stramm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinilello, Pamela Deneke
REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3349:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...426
SEQUENCE DESCRIPTION: SEQ ID NO: 3349:
US-09-107-433-3349

Query Match 5.7%; Score 103.5; DB 4; Length 426;
Best Local Similarity 20.8%; Pred. No. 0.017;
Matches 71; Conservative 54; Mismatches 131; Indels 85; Gaps 16;

10 DMHLHRLD-----GDVLKAVSHSAHFGRAIVMPNLKPPITTTAAVAAYREALIKS-- 61
61 DIHVHREBPQTHKEDIHTGALAAAGFTTVMMANTSPITSDVETLQAVLSAKKEI 120
62 -----LPVDSDFNPLMTLYITDFTTSPMEIKLARESOVFGVKLYPAGATTNSODGVTDLF 116
121 NVKTVATITGNFN---GKNLTDFKALLE-----AGAVGFSDDGIPLES 160
DB
117 GKCLPVLQGVENVHNMPLVHGEVTNPVDMFDEK-----VFIEVLRP----- 160
QY 161 SK---LVKEMAEAKKL-----NTFISLHEBDPGLNGVLGFENINAREHFIHICGATGV 210
DB 161 LRF-QGGLQPHNYCLPVLKREIHRBALVSATSGSKREFLTGDSAPH--DRRKECGCGC 263
QY 208 LTRQGSNAKKN---PRLRLSDRRAVIEGLKSGVIT-VIATDHAIPHVDKVEDITKAP 325
DB 264 AGIYMAFVALSV--YAKVFEKENALDKLEAFISFNGPDY 301
QY 326 SGMTGLTSLSLGLTYLVEAGELSLMELBKMT-YNPAKLY 365

RESULT 12
US-09-710-279-628
Sequence 628, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/09/710,279
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 628
LENGTH: 371
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-09-710-279-628

Query Match 5.7%; Score 102.5; DB 4; Length 371;
Best Local Similarity 22.7%; Pred. No. 0.018;

RESULT 13
US-09-710-279-260.

Query Match	5.7%	Score 102.5;	DB 4;	Length 425;
Best Local Similarity	22.7%;	Pred. No. 0.022;		

RESULT 14
US-09-134-001C-4803

Query Match	5.7%;	Score 102.5;	DB 3;	Length 427;
Best Local Similarity	22.7%;	Pred. No. 0.022;		
Matches 65;	Conservative 40;	Mismatches 114;		

RESULT 15
US-08-469

```

1 / Sequence:126, Application US/08469318
2 / Patent No. 6022535
3 /
4 / GENERAL INFORMATION:
5 / APPLICANT:
6 / TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion
7 / TITLE OF INVENTION: Protein
8 / NUMBER OF SEQUENCES: 196
9 / COMPUTER READABLE FORM:
10 / MEDIUM TYPE: Floppy disk
11 / COMPUTER: IBM PC Compatible
12 / OPERATING SYSTEM: PC-DOS/MS-DOS
13 / SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
14 / CURRENT APPLICATION DATA:
15 / APPLICATION NUMBER: US/08/469,318
16 / FILING DATE:
17 / CLASSIFICATION:
18 / PRIOR APPLICATION DATA:
19 / APPLICATION NUMBER: 08/446,872
20 / FILING DATE:
21 / INFORMATION FOR SEQ ID NO: 126:
22 / SEQUENCE CHARACTERISTICS:
23 / LENGTH: 244 amino acids
24 / TYPE: amino acid
25 / STRANDEDNESS:
26 / TOPOLOGY: linear
27 / MOLECULE TYPE: protein

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 05:13:50 ; Search time 92 Seconds

(without alignments)
1454.555 Million cell updates/sec

Title: US-10-070-277-2

Perfect score: 1812
Sequence: 1 MEISTIQPDMDHLRLRDGV.....SGDIIPFAGEMLDWLPAPL 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq_16Dec04:*
2: geneseqp1380s:*
3: geneseqp1390s:*
4: geneseqp2000s:*
5: geneseqp2001s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1812	100.0	346	4	AAB70773 Potato di
2	1812	100.0	346	4	AAB70779 Potato di
3	1528	84.3	345	5	AAO15493 Beta vulg
4	1067.5	58.9	364	6	ABU20245 Protein e
5	1058	58.4	352	6	ABU22502 Protein e
6	1039.5	57.4	350	6	ABU49750 Protein e
7	1008.5	55.7	354	6	ABU23462 Protein e
8	983	54.2	344	6	ABU37887 Protein e
9	979	54.0	344	6	ABP78043 N. gonorr
10	979	54.0	344	6	ABU37314 Protein e
11	978	54.0	348	6	ABU31804 Protein e
12	978	54.0	352	7	ABO61072 Klebsiell
13	975	53.8	348	6	ABU38565 Protein e
14	975	53.8	348	7	ABO68745 pseudomon
15	971	53.6	348	6	ABU15296 Protein e
16	962	53.1	348	6	ABU39671 Protein e
17	959	52.9	347	6	ABU42003 Protein e
18	956	52.8	348	6	ABU50480 Protein e
19	952.5	52.6	344	6	ABU16776 Protein e
20	952.5	52.6	346	6	ADA34254 Actinoba
21	949	52.4	348	6	ABU47954 Protein e
22	946	52.2	355	8	ADU04644 M. catarr
23	943	52.0	354	6	ABU35190 Protein e
24	935	51.6	351	6	ABM68113 Photornab
25	934	51.5	348	6	ABU44828 Protein e

26	930	51.3	356	7	ADF04000 Bacterial
27	559	31.4	335	6	ABU26343 Protein e
28	541	29.9	339	6	ABU30820 Protein e
29	539.5	29.8	190	6	ABU27626 Protein e
30	527	29.1	371	2	AAW20981 H. pylori
31	402	22.2	136	6	ABU40992 Protein e
32	350.5	19.3	116	5	ABP06300 Human ORF
33	209.5	11.6	146	2	AAW20558 H. pylori
34	144	7.9	396	6	ABU24485 Streptoco
35	134.5	7.4	438	5	ABP26430 Streptoco
36	131.5	7.3	426	5	ABBA47703 Listeria
37	131.5	7.3	426	6	ABU32919 Protein e
38	128	7.1	424	6	ABU24236 Protein e
39	126.5	7.0	409	8	ADNA47121 Thermococ
40	123	6.8	397	4	ADBI1184 Lactococc
41	121.5	6.7	427	4	AAU35265 Enterococ
42	121.5	6.7	427	6	ABU29149 Protein e
43	120	6.6	441	5	ABBS4385 Lactococc
44	119.5	6.6	426	6	ABU29702 Protein e
45	119.5	6.6	430	7	ADC95480 E. faeciu

ALIGNMENTS

RESULT 1
AAB70773 standard; protein; 346 AA.
ID AAB70773:
XX
AC AAB70773:
XX
DT 25-MAY-2001 (first entry)
XX
DE Potato dihydro-oxotase protein.
XX
KM Potato; dihydro-oxotase; plant; polysaccharide content; orotic acid;
KW uridine; starch synthesis.
XX
OS Solanum tuberosum.
XX
PN WO200114569-A2.
XX
PD 01-MAR-2001.
XX
PF 12-AUG-2000; 2000WO-EP007884.
XX
PR 20-AUG-1999; 99DE-01039688.
XX

(BADI) BASF AG.
Ehrhardt T, Stitt Nigel M, Geigenberger PL, Loeff I, Zrenner R;
Schroeder M;
WPI: 2001-202938/20.
DR N-PSDB; AAF61217.
XX
PT Increasing the polysaccharide, especially starch, content of plants, by
transforming with a sequence encoding dihydro-oxotase.
XX
PS Example 4; Page 30-31; 32pp; German.

This invention describes a novel use of a DNA sequence (I) encoding dihydro-oxotase (II) to prepare plants which have increased (1) a 1271 base polysaccharide (PS) content. The invention also describes (1) a 1271 base pair DNA sequence (I) encoding potato (II), sequences that hybridize to it and its homologs (or their fragments); (2) method for preparing plants of increased PS content by expressing (I) in them; (3) method for transforming plants by introducing an expression cassette, containing promoter, signal sequence and (I), into a plant cell, callus tissue, complete plant or protoplast; and (4) plants with increased PS content containing the DNA of (1). Expression of (II) increases the content of orotic acid (a precursor of uridine nucleotides), resulting in an increased rate of starch synthesis. Typically, in potato tuber slices,

CC addition of 10 mM orotate more than doubled the rate of starch synthesis.
 CC (1) is used to prepare plants (especially tomato, tobacco, potato,
 CC tapio, manioc, rice, barley, oats, rye, wheat and maize) with increased
 CC PS, particularly starch, content
 XX Sequence 346 AA;

Query Match 100.0%; Score 1812; DB 4; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1,6e-173;
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELSTTQDDMHLRLRGDYLKAVSHSAHFGRAIYMPKLPITTTAAVAAYREALIK 60
 Db 1 MELSTTQDDMHLRLRGDYLKAVSHSAHFGRAIYMPKLPITTTAAVAAYREALIK 60
 QY 61 SLVPDSDFNPLMTLYLDTTSPMEIKLARSQVVGKLYPAGATTNSODGVTDLFGKCL 120
 Db 61 SLVPDSDFNPLMTLYLDTTSPMEIKLARSQVVGKLYPAGATTNSODGVTDLFGKCL 120
 QY 121 PVLQEMVEHNMPLLVHGEVTNPEDVDFREKVFIEIYLRPLVQKPPOLKVMHEHTTTIDA 180
 Db 121 PVLQEMVEHNMPLLVHGEVTNPEDVDFREKVFIEIYLRPLVQKPPOLKVMHEHTTTIDA 180
 QY 181 VKFVESCTEGFVAATYTPQHLVLNNSLFGGLQPHNYCLPVLKREIHREALVSAVTSGS 240
 Db 181 VKFVESCTEGFVAATYTPQHLVLNNSLFGGLQPHNYCLPVLKREIHREALVSAVTSGS 240
 QY 241 KRFFLGTDSAPHDRRREKSCSCGAGIYNAPVALSYAVFEKENALDLEAFSTFNGPDF 300
 Db 241 KRFFLGTDSAPHDRRREKSCSCGAGIYNAPVALSYAVFEKENALDLEAFSTFNGPDF 300
 QY 301 YGLPRNNSKIKLSKTPMKVPBSFYASGDIIPMFAGEMLDWLPAPL 346
 Db 301 YGLPRNNSKIKLSKTPMKVPBSFYASGDIIPMFAGEMLDWLPAPL 346

RESULT 2
 AAB70779
 ID AAB70779 standard; protein; 346 AA.

AC AAB70779;
 XX
 DT 25-MAY-2001 (first entry)
 XX
 DE Potato dihydro-oroate protein.
 XX
 KM Potato; dihydro-oroate; inhibitor; herbicide; transgenic plant.
 XX
 OS Solanum tuberosum.
 XX
 PN WO200118190-A2.
 XX
 PD 15-MAR-2001.
 XX
 PF 02-SEP-2000; 2000MO-EP008581.
 XX
 PR 07-SEP-1999; 99DE-01042742.
 XX
 PA (BADI) BASF AG.
 XX
 PI Ehrhardt T, Lerchl J, Stilt Nigel M, Zrenner R, Schroeder M;
 XX
 DR WPI; 2001-235198/24.
 XX
 DR N-PSDB; AAF61262.
 XX
 PT New DNA encoding plant dihydrooroate, for producing herbicide-resistant
 PT plants and for screening for compounds that can inhibit dihydrooroate
 PT and that be used as herbicides.
 XX
 PS Claim 3; Page 31-32; 38pp; German.
 XX
 CC This invention describes a novel DNA sequence (1) containing the coding
 CC region for a plant dihydro-oroate (DHO), comprising a sequence (S1) of

CC 1271 base pairs (bp), given in the specification. The invention also
 CC describes (1) a DNA sequence (1a) that hybridizes to (S1), or its parts
 CC or derivatives produced by insertion, deletion or substitution, and
 CC encodes a protein with DHO activity; (2) a protein (11) containing at
 CC least 100 amino acids (aa) from a 346 aa sequence (S2), given in the
 CC specification; (3) identifying substances (111) that inhibit activity of
 CC plant DHO; (4) identifying substances (111a) which act as herbicides by
 CC inhibition of plant DHO; (5) a test system based on expression of (S1)
 CC for identifying herbicidal inhibitors of DHO; and (6) inhibitors of plant
 CC DHO identified by the system of (5). The products of the invention have
 CC herbicide activity. (1) is used: (i) for recombinant production of DHO in
 CC prokaryotic or eukaryotic cells; (ii) to produce a test system for
 CC identifying compounds that inhibit DHO, potentially useful as herbicides;
 CC and (iii) to produce transgenic plants that have increased resistance to
 CC DHO-inactivating herbicides. Recombinantly expressed DHO provides a
 CC reproducible and rapid (high throughput) screen for specific inhibitors
 CC of activity
 XX
 SQ Sequence 346 AA;

Query Match 100.0%; Score 1812; DB 4; Length 346;
 Best Local Similarity 100.0%; Pred. No. 1,6e-173;
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELSTTQDDMHLRLRGDYLKAVSHSAHFGRAIYMPKLPITTTAAVAAYREALIK 60
 Db 1 MELSTTQDDMHLRLRGDYLKAVSHSAHFGRAIYMPKLPITTTAAVAAYREALIK 60
 QY 61 SLVPDSDFNPLMTLYLDTTSPMEIKLARSQVVGKLYPAGATTNSODGVTDLFGKCL 120
 Db 61 SLVPDSDFNPLMTLYLDTTSPMEIKLARSQVVGKLYPAGATTNSODGVTDLFGKCL 120
 QY 121 PVLQEMVEHNMPLLVHGEVTNPEDVDFREKVFIEIYLRPLVQKPPOLKVMHEHTTTIDA 180
 Db 121 PVLQEMVEHNMPLLVHGEVTNPEDVDFREKVFIEIYLRPLVQKPPOLKVMHEHTTTIDA 180
 QY 181 VKFVESCTEGFVAATYTPQHLVLNNSLFGGLQPHNYCLPVLKREIHREALVSAVTSGS 240
 Db 181 VKFVESCTEGFVAATYTPQHLVLNNSLFGGLQPHNYCLPVLKREIHREALVSAVTSGS 240
 QY 241 KRFFLGTDSAPHDRRREKSCSCGAGIYNAPVALSYAVFEKENALDLEAFSTFNGPDF 300
 Db 241 KRFFLGTDSAPHDRRREKSCSCGAGIYNAPVALSYAVFEKENALDLEAFSTFNGPDF 300
 QY 301 YGLPRNNSKIKLSKTPMKVPBSFYASGDIIPMFAGEMLDWLPAPL 346
 Db 301 YGLPRNNSKIKLSKTPMKVPBSFYASGDIIPMFAGEMLDWLPAPL 346

RESULT 3
 AA015493
 ID AA015493 standard; protein; 345 AA.

AC AA015493;
 XX
 DT 03-OCT-2002 (first entry)
 XX
 DE Beta vulgaris dihydrooroate.
 XX
 KM Sugar beet; transgenic plant; osmotic stress tolerance;
 KM oxidative stress tolerance; casein kinase alpha catalytic subunit;
 KM dihydrooroate; translation initiation factor 1a; plant flowering;
 KM plant growth stimulation.
 XX
 OS Beta vulgaris.
 XX
 PN WO200252012-A2.
 XX
 PD 04-JUL-2002.
 XX
 PF 20-DEC-2001; 2001MO-EP015093.
 XX

PR 22-DEC-2000; 2000EP-00870319.
 PR 26-FEB-2001; 2001US-0271656P.
 XX
 PA (CROP-) CROPDDESIGN NV.
 PI Kanhonou RA, Serrano Salom R, Ros Palau R;
 XX WPI; 2002-583567/62.
 DR N-PSDB; AAL44159.
 XX
 PT Novel sugar beet genes involved in stress response, useful for rendering
 PT crops resistant to stress situations like osmotic stress caused by salt,
 XX drought, cold or frost thus increasing crop yield.
 PS
 XX Claim 28; Fig 6; 95pp; English.
 CC The invention comprises the amino acid and coding sequences of sugar beet
 CC (Beta vulgaris) genes that are involved in osmotic and oxidative stress
 CC tolerance in a plant. The Beta vulgaris genes of the invention encode the
 CC proteins: casein kinase alpha catalytic subunit; dihydroxylase;
 CC translocation initiation factor 1a; putative protein; and an unknown
 CC protein. The Beta vulgaris genes of the invention are useful for
 CC enhancing the osmotic and oxidative stress tolerance of a plant. The Beta
 CC vulgaris genes are also useful for stimulating plant growth. The casein
 CC kinase alpha catalytic subunit gene is useful for controlling the process
 CC of flowering in a plant. The present amino acid sequence represents the
 CC Beta vulgaris dihydroxylase
 XX
 SQ Sequence 345 AA;
 Query Match 84.3%; Score 1528; DB 5; Length 345;
 Best Local Similarity 82.0%; Pred. No. 6.9e-145;
 Matches 282; Conservative 34; Mismatches 28; Indels 0; Gaps 0;
 QY 1 MEISTGPPDMWHLTRDGVLLKAVVSHSAHFGRAIVMPNKKPITTTAAVAAYRAILK 60
 DB 1 MELTLTRPDMWHLTRDGVLLAAVAHPASARFPAIVMPNLRPVTTTGAIAIRKSI 60
 QY 61 SLVPDSDFNPMLTYLTDTTSPMEIKLARSQVFGVKLYPAGATTNSODGVTDLFGKCL 120
 DB 61 VLSPDSDFNPMLTYLTDTTSPMEIKLARSQVFGVKLYPAGATTNSODGVTDLFGKCL 120
 QY 121 PVLDGEMVHNMPLLVHGEVTNPEDVDFREKVFIEYLRVLVQKFPQLKVMMEHTTTDA 180
 DB 121 PVLEEMABQDMPLLVHGEVTNPEDVDFREKVFIEYLRVLVQKFPQLKVMMEHTTTDA 180
 QY 181 VKFVESCETGAVATVTPQHVLVNRNSLFOGGLQPHNYCPLVKREIHRALVSAVTS 240
 DB 181 VKFVESCETGAVATVTPQHVLVNRNSLFOGGLQPHNYCPLVKREIHRALVSAVTS 240
 QY 241 KRFLIGTDSAPHDRRRKSCGCGAGIYVAPALSVYAKFEKENALDKLEAFTSPNGPDF 300
 DB 241 KQYLGTDSDAPHRRRKSCGCGAGIYVAPALSVYAKFEKENALDKLEAFTSPNGPDF 300
 QY 301 YGLPRNNSKIKLSTKTPKVPESFSYASGDIIPMFAGEMLDWLP 344
 DB 301 YGLPRNNSKIKLSTKTPKVPESFSYASGDIIPMFAGEMLDWLP 344

PN MO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342823P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
 PI Wall D, Trawick JD, Carr GD, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA24115.
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 48169; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 364 AA;
 Query Match 58.9%; Score 1067.5; DB 6; Length 364;
 Best Local Similarity 60.4%; Pred. No. 1.9e-98;
 Matches 206; Conservative 37; Mismatches 95; Indels 3; Gaps 2;
 QY 3 LSTGPPDMWHLTRDGVLLKAVVSHSAHFGRAIVMPNKKPITTTAAVAAYRAILKSL 62
 DB 22 LSLARPPDMWHLTRDGVLLAVALPHTRARQFRAIVMPNKKPITTTAAQAAYREIILAL 81
 QY 63 PVDSDFNPMLTYLTDTTSPMEIKLARSQVFGVKLYPAGATTNSODGVTDLFGKCLPV 122
 DB 82 PAGMTFEPMLTYLTDTTSPMEIKLARSQVFGVKLYPAGATTNSODGVTDLFGKCLPV 140
 QY 123 LQENVENHMPLLVHGEVTNPEDVDFREKVFIEYLRVLVQKFPQLKVMMEHTTTDAVK 182

```

Db      141 LEAMQETGMPDLVHGSVTDASTIDLPDRKRVFIDRMVTEPLRDFGLKAVFEHITTKAD 200
QY      183 FVE--SCTEGFVAATVTPQHLYLNRSNLSFGGLQPHNYCLPVKREIHEALVSAVTSGS 240
Db      201 VYRDADAAPGLLGATITTAHLLYRNMLFVGCIIPHYCPLVLRERHVALVEAATSGN 260
QY      241 KRFFLGTDSPAPDRRKESCGCGAGIYNAPVALSVAKVEKNALDKLEAFNSGNDPF 300
Db      261 PRFFLGTDSPAHADAKETACGACGYTALHLELYAEADDTAGALDKLEGFASFGADF 320
QY      301 YGLPRNNSKIKLSKTPKVPESFSYASGDIIPMPAGMIDW 341
Db      321 YGLPRSAETVTLRREPWELPREIRPAETPVPLRGSETIWM 361

```

RESULT 5
ABU22502
ID ABU22502 standard; protein; 352 AA.
XX
XX ABU22502;
AC
XX
XX

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #8029.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Burkholderia mallei.

EN WO200277183-A2.

XX 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KI, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
XX WPI; 2003-029926/02.
DR N-PSDB; ACA26372.

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 50426; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid;
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 352 AA;

Query Match 58.4%; Score 1058; DB 6; Length 352;
Best Local Similarity 58.4%; Pred. No. 1.7e-97;
Matches 199; Conservative 45; Mismatches 95; Indels 2; Gaps 1;

```

QY      3 LSTIQPDWHLHLDGDLKAVSHSAHFGRAIWMNLKPPITTTAAVAVERAILKSL 62
Db      9 LTLARPDWHLHVRDGAWLAAVLPHTARQGRBALIMNLPVTTTAAQAYERILAAV 68
QY      63 PVDSDPNPLMTLYLTPTTSPEWETKLARESOVFGVKLYPAGATTNSQDVTDLFGKCLPV 122
Db      69 PAGMTFELMTLYLTDTTPADEIRARESGCVGKLYPAGATTNSDAGVTDLLGKCAIT 128
QY      123 LOENVEHNPILLVGEVTNPEVDMFDEKVFIEFTVLRPIVOKRPOLKVMHEVTTIDAVK 182
Db      129 LEAMQEVGMPPLVHGEVTDPSIDLPRKEKVIDRWVEPLRRALPGLVFEHITTKAD 188
QY      183 FVE--SCTEGFVAATVTPQHLYLNRSNLSFGGLQPHNYCLPVKREIHEALVSAVTSGS 240
Db      189 VYRDADAASRGATITTAHLLYRNMLFVGCIIPHYCPLVLRERHVALVEAATSGN 248
QY      241 KRFFLGTDSPAPDRRKESCGCGAGIYNAPVALSVAKVEKNALDKLEAFNSGNDPF 300
Db      249 PRFFLGTDSPAHADAKETACGACGYTALHLELYAEADDTAGALDKLEGFASFGADF 308
QY      301 YGLPRNNSKIKLSKTPKVPESFSYASGDIIPMPAGMIDW 341
Db      309 YGLPRSAETVTLRREPWELPREIRPAETPVPLRGSETIWM 349

```

RESULT 6
ABU49750
ID ABU49750 standard; protein; 350 AA.
XX
XX ABU49750;
AC
XX
XX

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #35277.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Vibrio cholerae.

EN WO200277183-A2.

XX 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI: 2003-029926/02.
DR N-PSDB; ACA53620.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 77674; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
CC
SQ Sequence 350 AA;
Query Match 57.4%; Score 1039.5; DB 6; Length 350;
Best Local Similarity 59.0%; Pred. No. 1.2e-95;
Matches 200; Conservative 45; Mismatches 91; Indels 3; Gaps 3;
OY 3 LSTGPPDDMHNLKRGDVLKAVNSAHHPGATVMPULKPPITTAANAAYREALTSL 62
DB 12 LTTTRDDMHNLKRGDVLADTVDRDISKNGALMPTVPBVTTEALMAYREIMAAQ 71
OY 63 FVDSDFNPLMTLYLTDTTSPMEIKLARBSQVFFGVKLYPAGATTNSQDGVTLDFCKLPLV 122
DB 72 P-QAFPEFLMALYLTDNTSPERIRAKXSGKVAAKLVPAGATTNSDGSQVISA-KNIVPV 129
OY 123 LOEVENHNPMLLVHSEVTNPEVDMFDEKREVFETVLRPLVQKFPOLKVMMEHTTIDAVK 182
DB 130 LQAMQEVGMILLVHSEVTTHEDVIDFDRKXFTDVLAVIYVDFPOLKTVLBHITTDADV 189
OY 133 FVESTGESFVATVTPHULVILNBSLFGGLOPHNYCPVLKRETHREALVAVTSQSKR 242
DB 190 FVQQAQDN-VAAFTTAHLLFNRMMLVGGIRPHFYCPILKRAHGHALVAATSSGSK 248
OY 243 FFLGTDSAPHDRRRKSCGCGAGIYNAVALSVAVKEKENALDKLEATFSFNGPDFYG 302
DB 249 FFLGTDSAPHAKGREACGCGAGSYTAHAALVAVKEKEKLEALFAFNGPDFYG 308
OY 303 LPRNNSKITLSTKTPMKVPSFSYAGDITPMFAGMLDW 341

DB 309 LPRNGETVTLTKQAMPVAESMPGSDIVVPIRAGENIEW 347
RESULT 7
ABU23462
ID ABU23462 standard; protein; 354 AA.
XX
AC ABU23462;
XX
PT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #9989.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS *Bordetella pertussis*.
XX
PN MO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI: 2003-029926/02.
DR N-PSDB; ACA27332.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 51386; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 354 AA;

Query Match 55.7%; Score 1008.5; DB 6; Length 354;
Best Local Similarity 57.0%; Pred. No. 1.6e-92;
Matches 196; Conservative 46; Mismatches 97; Indels 5; Gaps 3;

```
QY 2 ELSTQDDPDMHLNRDGDVLAQVSHSAHFGRALVWPNKXPITTTAAVAVR--EAI 58
Db 8 ELTTRPDDMHLNRDGSALAEAVLDTRAPALVWPNRPVVTTEQALVYRARIENA 67
QY 59 LKSPVQ-SDPNPMNTLYTDTTSPMEIKLARSQVFGVKLYPAGATTNSODGVTDLEFG 117
Db 68 LKRGSGTAAFTPLMTLYTDTNPABEIVAHESGQVAAVKLYPAGATTNSAGTTDLG 127
QY 118 KCLPVLQEWVENHNPPLVHGEVNPPEVDMREKVEIEVLAFLVQKFPOLKVMEHVT 177
Db 128 KCGAALALAEKRCGMPPLVHGEVTDPAIDVFDEAVFIERVWQPLRRAPGLKVFEHITT 187
QY 178 IDAVKPFVSECTEGFVAATVTPQHLVINRNSLFOGGLQPHNYCLPVLEKEIHREALVSAVT 237
Db 188 REGAHYVRD-ASGPTATTTPOHMLYNRNAIFTGVRPHWCLPILKEVHRQLVEMAT 246
QY 238 SSGSKPEFLGTDNAPHRRRKSCSCGAGIYNAPVALSVYAKVFEKENALDKLEAPTSFNG 297
Db 247 SSGSPRFFLGTDNAPHRGKKEHACCGACGCTTALHAMELYATAFAVGRLDLKLEAFSGFHG 306
QY 298 PDYVGLPRNNSKIKLSKTPWKPESFSYASGDIIPMFAGEMLDW 341
Db 307 PDYVGLPRNNTGTLTLRREAVETPAEVAFGDTTLVPLSGSESLGW 350
```

RESULT 8
ABU37887
ID ABU37887 standard; protein; 344 AA.
XX
AC ABU37887;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #23414.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Neisseria meningitidis.
XX
BN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX MPI; 2003-029926/02.
XX
DR N-PSDB; ACA41757.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 65811; 1766pp; English.
XX

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC in electronic form part of the printed specification, but was obtained
CC from ftp.wipo.int/pub/published_pct_sequences

Sequence 344 AA;

Query Match 54.2%; Score 983; DB 6; Length 344;
Best Local Similarity 56.3%; Pred. No. 5.8e-90;
Matches 192; Conservative 55; Mismatches 88; Indels 6; Gaps 5;

```
QY 3 LSTQDDPDMHLNRDGDVLAQVSHSAHFGRALVWPNKXPITTTAAVAVREALKSL 62
Db 4 LTTTRPDDMHLNRDSDALKAAVAPYARQMGRAVIMPNLKPVSVAADALVAKRIMAAAL 63
QY 63 PVSDPNPMNTLYTDTTSPMEIKLARSQVFGVKLYPAGATTNSODGVTDLEKCLPV 122
Db 64 PEGSAPEPLMTLYLTQATPALVREKAAQIV-AKLYPAGATTNSDGVTDLE-KLIPV 121
QY 123 LQEMVENHNPPLVHGEVNPPEVDMREKVEIEVLAFLVQKFPOLKVMEHVTITDAVK 182
Db 122 LEMNAKQGLFVHGEVTDPEIDIPDEAAFIQVMEKVLQAVPNLKVFEHITTAAR 181
QY 183 FVESCTEGFVAATVTPQHLVINRNSLFOGGLQPHNYCLPVLEKEIHREALVSAVTSSK- 241
Db 182 LVLEAGDN-VAATVTPQHLVINRNDLVGVGRPHFCPLVLEKETHQALVAAT-GEKA 239
QY 242 -RFFLGTSAPHRDRKSCSCGAGIYNAPVALSVYAKVFEKENALDKLEAFSTFNGPDF 300
Db 240 HKFFLGTSAPHRDRKSCSCGAGIYNAPVALSVYAKVFEKENALDKLEAFSTFNGPDF 299
QY 301 YGLPRNNSKIKLSKTPWKPESFSYASGDIIPMFAGEMLDW 341
Db 300 YGIPENTDTITLVKQSGTVPASISYGDGLVPMRAGGEIGW 340
```

RESULT 9
ABP78043
ID ABP78043 standard; protein; 344 AA.
XX
AC ABP78043;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae amino acid sequence SEQ ID 2616.
XX

KW Antidiarrheal, infection, vaccine, gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN MO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB002069.
XX
PR 12-FEB-2001; 2001GB-00003424.
XX
PI (CHIR-) CHIRON SPA.
XX Fontana MR, Piazza M, Maignani V, Monaci E;
PI WPI; 2003-058415/05.
XX N-PSDB; ABZ39013.
XX
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection.
PS Disclosure; Page 379; 815pp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX
SQ Sequence 344 AA;

Query Match 54.0%; Score 979; DB 6; Length 344;
Best Local Similarity 56.3%; Pred. No. 1.5e-89;
Matches 192; Conservative 55; Mismatches 88; Indels 6; Gaps 5;

QY 3 LSIQPDWMHLHLDGDLKAVVSHAHFGRALVMPNLPKPPITTTAAVAYRAILKSL 62
DB 4 LTIIRPDWMHLHLDGDLKAVVSHAHFGRALVMPNLPKPPITTTAAVAYRAILKSL 63
QY 63 PVDSDPMLTLVLTDTTSPEIKLARESQVVFVKVYPAGATTNSODGVTLDFGKCLPV 122
DB 64 PEGSAFERLMTLYLTLDQATPELVBAKAGIV-AFKLYPAGATTNSDGVTLDF-KLIPV 121
QY 123 LQENVEHNMPLLVHGEVTPVDMPDREKVIETVLRPLVOKFQOLKVMHEHTTIDAVK 182
DB 122 LEEMAKQGIPLFVHGEVTPDEIDIFDREAFAFGVMKPVLAQVPLKVFPHITTAEMAR 181
QY 183 FVSECTBEFVAATVPPQHLVYNRNSLFCGGIOPHNVCPLVKRIHREALVSAATSSGSK- 241
DB 182 LVLEAGDN-VAATVTPQHLNRLNDLVGVRPHNFCPLVKRETHRALVAATV-GEKA 239
QY 242 -RFELGTDSAPDRRRKESCGCAGIYNAVPALSVYAVAFKEXNALDCLBAFTSNGPDF 300
DB 240 HKFPLGDSAPHAHSAKKNACGCMGFAMPAHIELYAVFEKAGLDTLEAFASNGGARF 299
QY 301 YGLPRNNSKITLSTKTPWVPSFSYASGDIIPMFGEMLDW 341
DB 300 YGIPENADTITLVKOSQTVPAVSPYGDGLVPMRAGEIGW 340

RESULT 10
ABU37314
ID ABU37314 standard; protein; 344 AA.
XX
AC ABU37314;
XX
DT 23-OCT-2003 (revised)
DT 19-JUN-2003 (first entry)
XX

DE Protein encoded by Prokaryotic essential gene #22841.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Neisseria gonorrhoeae.
XX
PN MO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (BLIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA41184.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
PS Claim 25; SEQ ID NO 65238; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC parent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 344 AA;

Query Match 54.0%; Score 979; DB 6; Length 344;
Best Local Similarity 56.3%; Pred. No. 1.5e-89;
Matches 192; Conservative 55; Mismatches 88; Indels 6; Gaps 5;

QY 3 LSIQPDWMHLHLDGDLKAVVSHAHFGRALVMPNLPKPPITTTAAVAYRAILKSL 62
DB 4 LTIIRPDWMHLHLDGDLKAVVSHAHFGRALVMPNLPKPPITTTAAVAYRAILKSL 63
QY 63 PVDSDPMLTLVLTDTTSPEIKLARESQVVFVKVYPAGATTNSODGVTLDFGKCLPV 122
DB 64 PEGSAFERLMTLYLTLDQATPELVBAKAGIV-AFKLYPAGATTNSDGVTLDF-KLIPV 121
QY 123 LQENVEHNMPLLVHGEVTPVDMPDREKVIETVLRPLVOKFQOLKVMHEHTTIDAVK 182
DB 122 LEEMAKQGIPLFVHGEVTPDEIDIFDREAFAFGVMKPVLAQVPLKVFPHITTAEMAR 181
QY 183 FVSECTBEFVAATVPPQHLVYNRNSLFCGGIOPHNVCPLVKRIHREALVSAATSSGSK- 241
DB 182 LVLEAGDN-VAATVTPQHLNRLNDLVGVRPHNFCPLVKRETHRALVAATV-GEKA 239
QY 242 -RFELGTDSAPDRRRKESCGCAGIYNAVPALSVYAVAFKEXNALDCLBAFTSNGPDF 300
DB 240 HKFPLGDSAPHAHSAKKNACGCMGFAMPAHIELYAVFEKAGLDTLEAFASNGGARF 299
QY 301 YGLPRNNSKITLSTKTPWVPSFSYASGDIIPMFGEMLDW 341
DB 300 YGIPENADTITLVKOSQTVPAVSPYGDGLVPMRAGEIGW 340

RESULT 10
ABU37314
ID ABU37314 standard; protein; 344 AA.
XX
AC ABU37314;
XX
DT 23-OCT-2003 (revised)
DT 19-JUN-2003 (first entry)
XX

Db 4 LTIIRPDDMHILNRDGDALKAVALPYARQWGRAVIMENLKPVVSVADALAKATIMAL 63
QY 63 PVSDPFPNLTLYLTDTTSPMEIKLARESQVFGVKLYPAGATTNSQDGVTLDFGKCLPV 122
Db 64 PESSAPFPLMTLYLTQATPELVREKKAQIV-AFLCLYPAGATTNSDGVTLDF-KLIPV 121
QY 123 LQEMVENHNPFLVHGEVTNPEVDMFDRKVFIEIVLRPVQKFPOLKVMHEHTIDAVK 182
Db 122 LERMAKQGLFLVHGEVTDPEIDIFDREAFFIGRWKPVLAQVPLKVFHEHTIDAVK 181
QY 183 FVESCIEGFVAATVTPPOHVLVNRNSLFQGLQPHNYCLPVLRKREIHREALVSAVTSGRK 241
Db 182 LVLEAGDN-VAAIVVTPHQLLNRLNDLVGVAPRHHFCLPVLRKRETHQALVAATV-GRKA 239
QY 242 -RFFICTDSAPDRRRKSCSCGAGIYNAPVALSYAVKFEKENALDKLEAFTSPNGPFD 300
Db 240 HKEFICTDSAPRAKSKAKENACGACGAFSMTHAIEIYAEVFEKAGLIDKLEAFASKNGARF 299
QY 301 YGLPNNNSKIKLSTKTPMKVPESFSYASGDIIPMFAGEMLDW 341
Db 300 YGIPENADTTTLVQSQTVPAVSPYGDGLVPMRAGEIGW 340

RESULT 11

ABU31804

ID ABU31804 standard; protein; 348 AA.

XX AC ABU31804;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #17331.

XX KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Klebsiella pneumoniae.

XX PN W020027183-A2.

XX PD 03-OCT-2002.

XX PE 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (EITR-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KJ, Zyckind JW,

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA35674.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

XX PT PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 59728; 1766bp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

XX CC the 6213 antisense sequences given in the specification where expression

XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX CC encoding a polypeptide whose expression is inhibited by the antisense

XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX CC polypeptide or its fragment whose expression is inhibited by the

XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC product is overexpressed or underexpressed; (11) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

XX Sequence 348 AA;

XX SQ

Query Match 54.0%; Score 978; DB 6; Length 348;

Best Local Similarity 55.8%; Pred. No. 1.9e-89;

Matches 189; Conservative 46; Mismatches 102; Indels 2; Gaps 2;

QY 3 LTIIRPDDMHILNRDGDALKAVALPYARQWGRAVIMENLKPVVSVADALAKATIMAL 62

Db 8 LKIRPDDMHILNRDGDALKAVALPYARQWGRAVIMENLKPVVSVADALAKATIMAL 67

QY 63 PVSDPFPNLTLYLTDTTSPMEIKLARESQVFGVKLYPAGATTNSQDGVTLDFGKCLPV 122

Db 68 PAGESFPNLTLYLTDTTSPMEIKLARESQVFGVKLYPAGATTNSQDGVTLDFGKCLPV 126

QY 123 LQEMVENHNPFLVHGEVTNPEVDMFDRKVFIEIVLRPVQKFPOLKVMHEHTIDAVK 182

Db 127 LERMAKQGLFLVHGEVTDPEIDIFDREAFFIGRWKPVLAQVPLKVFHEHTIDAVK 186

QY 183 FVESCIEGFVAATVTPPOHVLVNRNSLFQGLQPHNYCLPVLRKREIHREALVSAVTSGRK 242

Db 187 YVRDGNB-LIAAATTPHQLLNRLNDLVGVAPRHHFCLPVLRKRETHQALVAATV-GRKA 245

QY 243 FFLICTDSAPDRRRKSCSCGAGIYNAPVALSYAVKFEKENALDKLEAFTSPNGPFD 302

Db 246 AFLICTDSAPRAKSKAKENACGACGAFSMTHAIEIYAEVFEKAGLIDKLEAFASKNGARF 305

QY 301 YGLPNNNSKIKLSTKTPMKVPESFSYASGDIIPMFAGEMLDW 341

Db 306 LPVIESYVELVRETTVVDISALPNDTLVPLAGETVRW 344

RESULT 12

ID ABO61072 standard; protein; 352 AA.

XX AC ABO61072;

XX DT 29-JUL-2004 (first entry)

XX DE Klebsiella pneumoniae polypeptide seqid 7589.

XX KM Recombinant expression vector; transcription regulatory element;

XX KM Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX OS Klebsiella pneumoniae.

XX PN US6610836-B1.

XX PD 26-AUG-2003.

XX PF 27-JAN-2000; 2000US-00489039.

Db 184 FVEAPAN-VGATITAHLLYNRNHMLVGGIRPHFYCLPIKRNTHOEALLDAAVSGNPK 242
QY 243 FFLGTDAPHDHRRKESCGCAGIYNAVALSVYAKVFEKENALDKLEAFSTFNGDPFYG 302
Db 243 FFLGTDAPHDHRRKESCGCAGIYNAVALSVYAKVFEKENALDKLEAFSTFNGDPFYG 302
QY 303 LPRNNSKIKLSTKTPMKVPESESFYASGDIIPMFAGEMLDW 341
Db 303 LPRNTRITLTVREEMQAPASLPFGDPDVPLRAGETLRW 341

RESULT 14

ABO68745
ID ABO68745 standard; protein; 398 AA.
XX ABO68745;
AC ABO68745;
XX 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #920.
XX
KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
DR N-PSDB; ABD02316.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 17491; 455bp; English.
XX

The invention relates to Pseudomonas aeruginosa polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics, bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a P. aeruginosa nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs, including anti-P. aeruginosa drugs, as templates for recombinant production of P. aeruginosa-derived peptides or polypeptides, as target components for diagnosis and/or treatment of P. aeruginosa-caused infection, and in detection of P. aeruginosa sequences or other sequences of Pseudomonas species using bioclip technology. Sequences ABO67826-CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html

XX
SQ Sequence 398 AA;

Query Match 53.8%; Score 975; DB 7; Length 398;
Best Local Similarity 55.5%; Pred NO.4.7e-89;
Matches 188; Conservative 43; Mismatches 106; Indels 2; Gaps 2;

QY 3 LSTITPDDWHLRLRGDVLKAVSHSAHFRATVMPKLPITTTAAVAYRAITKSL 62
Db 55 LTLRPDDWHLRLRGDVLKAVSHSAHFRATVMPKLPITTTAAVAYRAITKSL 62
55 LTLRPDDWHLRLRGDVLKAVSHSAHFRATVMPKLPITTTAAVAYRAITKSL 62
114

QY 63 PVSDRNPMLTLYLDTTSPMEIKLARESOVVGKLYPAGATTNSODGVTDLPFKCLPV 122
Db 115 PAASREPLMLVLTLTERTSAEIRPTAKASGFYAAKLYPAGATTNSDGVTRI -DNIEA 173
QY 123 LQEMVHNHNPMLVHGVTNPEVDMFDRKVFETLRLVQKFPOLKVMHEVTTIDAVK 182
Db 174 LEMAVGNPMLVHGVTNPEVDMFDRKVFETLRLVQKFPOLKVMHEVTTIDAVK 233
QY 183 FVESCTEGVAAVTTPQHLVLRNSLFOGQLOPHNYCLVLRKREHREALVSAVTSQSKR 242
Db 234 FVEAPAN-VGATITAHLLYNRNHMLVGGIRPHFYCLPIKRNTHOEALLDAAVSGNPK 292
QY 243 FFLGTDAPHDHRRKESCGCAGIYNAVALSVYAKVFEKENALDKLEAFSTFNGDPFYG 302
Db 293 FFLGTDAPHDHRRKESCGCAGIYNAVALSVYAKVFEKENALDKLEAFSTFNGDPFYG 352
QY 303 LPRNNSKIKLSTKTPMKVPESESFYASGDIIPMFAGEMLDW 341
Db 353 LPRNTRITLTVREEMQAPASLPFGDPDVPLRAGETLRW 391

RESULT 15

ABU15296
ID ABU15296 standard; protein; 348 AA.
XX
AC ABU15296;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #823.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Escherichia coli.
XX
PN WO20027183-A2.
XX
PD 03-OCT-2002.
XX

PF 21-MAR-2002; 2002WC-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR N-PSDB; ACA19166.
XX

PT New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX

PS Claim 25; SEQ ID NO 43220; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 348 AA:

Query Match 53.6%; Score 971; DB 6; Length 348;

Best Local Similarity 55.8%; Pred. No. 9.6e-89; Mismatches 102; Indels 2; Gaps 2;

Matches 189; Conservative 46; Mismatches 102; Indels 2; Gaps 2;

```
QY 3 LSTPPDDMHLMRDGDLKAVVSHSAHFGRALVMPNLKPITTTAAVAAYREALIKSL 62
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8 LKIRPDDMHLMRDGDLKAVVSHSAHFGRALVMPNLKPITTTAAVAAYREALIKSL 67
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 63 PVDSDFNPLMTLVLTDTTSPMEIKLARESQVFGVKLYPAGATTNSQDGVTDLFGKCLPV 122
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 68 PAGHDFPLMTLVLTDTTSPMEIKLARESQVFGVKLYPAGATTNSQDGVTDLFGKCLPV 126
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 123 LQENVEHNMPLVHGEVTNPEVDMFDEKXFIETVLRPLVQKPPQKLVNMEHTTIDAVK 182
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 127 LERMEKIGMPLVHGEVTNPEVDMFDEKXFIETVLRPLVQKPPQKLVNMEHTTIDAVK 186
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 183 FVSGCTEGFVAATVTPQHLVLRNSLFOGGLQPHNYCLPYLKRRIHREALVSAVTSQSKR 242
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 187 YVRDGNR-LAATTPQHLVLRNSLFOGGLQPHNYCLPYLKRRIHREALVSAVTSQSKR 245
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 243 FFLGTDSAPHDRREKSCSCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFGPDYFG 302
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 VFLGTDSAPHDRREKSCSCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFGPDYFG 305
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 303 LPRNNSKIKLSKTPWKVPESFSYASGDIIPMFAGEMLDW 341
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 306 LPVNDTFIELVREQVAESIALTDTLVPLAGETVRW 344
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Search completed: February 12, 2005, 07:02:59
Job time : 96 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2005, 05:25:37 ; Search time 102 Seconds
(without alignments)
1737.053 Million cell updates/sec

Title: US-10-070-277-2

Perfect score: 1812
Sequence: 1 MELSTIQDDMHLLHRDGV.....SGDIIMFAGEMLDMLPAPL 346

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1528	84.3	345	2 Q8VXS9	Q8VXS9 beta vulgar
2	1467	81.0	377	1 PYRC_ARATH	O04904 arabidopsis
3	1458	80.5	359	2 Q94J38	Q94J38 oryza sativ
4	1083	59.8	349	2 Q82WF3	Q82WF3 nitrosoma
5	1061	58.6	344	1 PYRC_RALSO	O8Y249 ralsoma s
6	1058	58.4	352	2 Q62HT5	Q62HT5 burkholderi
7	1058	58.4	352	2 Q63QW0	Q63QW0 burkholderi
8	1056.5	58.3	342	1 PYRC_SYNY3	P74438 synchocyst
9	1053	58.1	361	2 Q7NQ12	Q7NQ12 chromobacte
10	1039.5	57.4	342	1 PYRC_VIBCH	Q9K124 vibrio chol
11	1012.5	55.9	342	2 Q6LPI7	Q6LPI7 photobacter
12	1012.5	55.9	359	2 Q7M4A4	Q7M4A4 bordetella
13	1012.5	55.9	359	2 Q7WFR2	Q7WFR2 bordetella
14	1008.5	55.7	359	2 Q7VUP3	Q7VUP3 bordetella
15	1000	55.2	343	2 Q8EB40	Q8EB40 shewanella
16	991.5	54.7	342	1 PYRC_VIBPA	O8J146 vibrio para
17	986.5	54.4	347	2 Q7MFR3	Q7MFR3 vibrio vuln
18	983	54.2	344	1 PYRC_NEIMA	O9YV66 neisseria m
19	982	54.2	344	1 PYRC_NEIMA	O9K0D1 neisseria m
20	979.5	54.1	342	1 PYRC_VIBVU	O83K19 vibrio vuln
21	975	53.8	348	2 PYRC_PSEAB	P72170 pseudomonas
22	972	53.6	348	2 Q83RT8	Q83RT8 shigella fi
23	971	53.6	347	1 PYRC_ECOLI	P05020 escherichia
24	970	53.5	347	1 PYRC_ECO57	O8X8N8 escherichia
25	967	53.4	347	2 Q6D658	Q6D658 erwinia car
26	966	53.3	347	1 PYRC_ECOL6	O8H172 escherichia
27	962	53.1	348	1 PYRC_PSEBK	O8BNW7 pseudomonas
28	959	52.9	347	1 PYRC_PSEBM	Q87XN1 pseudomonas
29	955	52.8	348	1 PYRC_YERPE	O8FNU4 yersinia pe
30	955	52.7	348	2 Q669R2	O669R2 yersinia ps
31	951	52.5	345	1 PYRC_AGR15	O8U199 agrobacteri

32	951	52.5	347	1 PYRC_SALTY	P06204 salmonella
33	949	52.4	347	1 PYRC_SALTI	O8Z712 salmonella
34	939.5	51.8	344	2 Q6PD29	O6FD29 acinetobact
35	935	51.6	350	2 Q7NSW0	Q7NSW0 photorhabdu
36	931	51.4	347	1 PYRC_RHIME	Q928C7 thizobium m
37	925	51.0	351	2 Q83B17	Q83B17 coxiella bu
38	913	50.4	343	2 Q7V618	Q7V618 synchococc
39	901	49.7	349	2 Q7V2B2	Q7V2B2 prochloroco
40	899.5	49.6	359	2 Q7VD18	Q7VD18 prochloroco
41	893.5	49.3	354	2 Q7W896	Q7W896 prochloroco
42	859.5	47.4	350	1 PYRC_BUCAI	P57416 buchiera ap
43	856.5	47.3	349	1 PYRC_BUCAP	Q8K912 buchiera ap
44	697.5	38.5	349	2 Q8D3B4	Q8D3B4 wigglewort
45	596.5	32.9	204	1 PYRC_SERMA	Q9S361 serratia ma

ALIGNMENTS

RESULT 1	ID	Q8VXS9	PRELIMINARY	PRT	345 AA.
AC	Q8VXS9				
DT	01-MAR-2002 (T-EMBLrel. 20, Created)				
DT	01-MAR-2002 (T-EMBLrel. 20, Last sequence update)				
DT	01-OCT-2003 (T-EMBLrel. 25, Last annotation update)				
DE	Dihydroorotase.				
GN	Name=dhase;				
OS	Beta vulgaris (Sugar beet).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
OC	Caryophyllales; Amaranthaceae; Beta.				
OX	NCBI_TaxID=161934;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Kanounou R., Serrano R., Ros R.;				
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.				
CC	-1- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-				
CC	aspartate.				
CC	-1- COFACTOR: Binds 2 zinc ions per subunit (By similarity).				
CC	-1- PATHWAY: Pyrimidine biosynthesis; third step.				
CC	-1- SIMILARITY: Belongs to the DHOase family.				
DR	EMBL; AJ251897; CAC80990.1; -.				
DR	HSSP; P05020; 1079				
DR	GO; GO:0004151; F:dihydroorotase activity; IEA.				
DR	GO; GO:0016787; F:hydrolyase activity; IEA.				
DR	GO; GO:0019856; P:pyrimidine base biosynthesis; IEA.				
DR	GO; GO:0006221; P:pyrimidine nucleotide biosynthesis; IEA.				
DR	Pfam; PF01979; Amidohydro 1; 1.				
DR	TIGRFam; TIGR00856; PYRC_dimer; 1.				
DR	PROSITE; PS00482; DIHYDROOROTASE_1; 1.				
DR	PROSITE; PS00483; DIHYDROOROTASE_2; 1.				
KW	Hydroxylase; Metal-binding; Pyrimidine biosynthesis; zinc.				
SQ	SEQUENCE 345 AA; 38297 MW; 8CA937E20895763F CRC64;				
Query Match	84.3%;	Score 1528;	DB 2;	Length 345;	
Best Local Similarity	82.0%;	Pred. No. 8.8e-116;			
Matches 282;	Conservative 34;	Mismatches 28;	Indels 0;	Gaps 0;	
Qy	1	MELSTIQDDMHLLHRDGVLLKAVSHSAHFGRAIVMPNLKPIITTTAAVAAYREALK	60		
Db	1	MELTUTRPDDMHLLHRDGLLAVAAPSAHFGRAIVMPNLKPIITTTGAIAARKIME	60		
Qy	61	SLPVSDFNPMLTYLTDTTSPMEIKLARSQVFGVKLYPAGATTNSQDGVTLFGKCL	120		
Db	61	VLPPDSDFNPMLTYLTDTTSPNEIKLARSEVVYAVKLYPAGATTNSQDGVTLFGKCL	120		
Qy	121	PVLQEMFENMPLLVHGEVTPNPEVMDREKVFETLRLPVLQKFMVEHTTIDA	180		
Db	121	PVLQEMFQDMPLLVHGEVTPDVIDFIREKVFETLRLPVLQKFMVEHTTIDA	180		
Qy	181	VKPFESCTEGFVAATVPQHLVLRNSLFOGGLQPHNYCLPVLKREIHTREALVSAVTSGS	240		

DB 181 VKFIESGNGVATVTPQHVLVNRNSLFGGGLQPHNYCLPULKEIRHQAIVSAVTS 240
 QY 241 KRFLGTDSAPHRBRRECCGAGIYNAPVALSVYAKVFEKENALDKLEAFSTFNGPDF 300
 DB 241 KQYFLGTDSAPHERREKSCGAGIYNBPVALSLYAKVFEBAALDKLEAFSTFNGPDF 300
 QY 301 YGLPRNNSKITLSTKTPKVPESFSYASGDIIPMFAGEMDMLP 344
 DB 301 YGLPRNNSKITLSTKTPKVPESFSYASGDIIPMFAGEMDMLP 344
 RESULT 2
 ID PYRC_ARATH STANDARD; PRT; 377 AA.
 AC 004904;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Dihydroorotase, mitochondrial precursor (EC 3.5.2.3) (DHOase).
 GN Name:PYR4; OrderedLocustNames=At4g22930; ORFNames=F7H19.110;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 [1]
 [2]
 SEQUENCE FROM N.A.
 RC STRAIN=CV. Columbia;
 RA Zhou L., Lacroute F., Thornburg R.W.;
 RT "Characterization of the Arabidopsis thaliana cDNA encoding
 (ex) Plant Gene Register PR97-115.
 RN [2]
 RC STRAIN=CV. Columbia;
 RA MEDLINE=20083488; PubMed=10617198; DOI=10.1038/47134;
 RA Mayer K.F.X., Schueller C., Wambolt R., Murphy G., Volckaert G.,
 RA Poll T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Terry N.,
 RA Weichselgartner M., de Simone V., Obermaier R., Maché R., Mueller M.,
 RA Kreis M., Delaeny M., Puidomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Porteleille D., Perez-Alonso M., Bouvier M., Bancroft I.,
 RA Vos P., Hohenseil J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham J., Robben J.,
 RA Van der Schueren J., Gymnopoulos B., Chuang Y.-J., Vandebussche F.,
 RA Brecken M., Weljens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Meitznerger T., Bothe G., Kampsperger U., Hilbert H., Braum M.,
 RA Holzner E., Brandt A., Peters S., van Staveren M., Dirke W.,
 RA Moolenaar P., Klein Lankhorst R., Rose M., Hauf J., Koetler P.,
 RA Berneriser S., Hempel S., Feldpausch M., Lambrecht S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin J., Quail M.A., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McClay K., Mayes R.,
 RA Pettett A., Rajadream M.A., Lyne M., Benes V., Reckmann S.,
 RA Borkova D., Bloecher H., Scharte M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel S., Fuchs M., Faltmann B., Grandert K., Danner D., Herzl A.,
 RA Neumann S., Argitlou A., Vitale D., Lignori R., Piravandi E.,
 RA Massenet O., Quigley F., Clabaud G., Muehlenstein A., Felber R.,
 RA Schenab S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Cheffor F., Cooke R., Bergeret C., Montfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Furnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Frisman D., Haase D., Lemcke R., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedha N., Gnoj L., Schütz R., Huang B., Spielgeid L.,
 RA Stoeckling P., Kalicki J., Graves T., Cordes M., Abu-Threideh J.,
 RA Larcelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Krumer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
 RA Nelson J., Speith J., Ryan E., Andrews S., Geisler C., Layman D.,

RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidan M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vill D., Sheher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Granat S., Shohdy N., Haegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Maira M.A., Marijansen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:769-777(1999).
 CC -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-
 aspartate.
 CC -1- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
 CC -1- PATHWAY: Pyrimidine biosynthesis; third step.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (Potential).
 CC -1- SIMILARITY: Belongs to the DHOase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF00146; AAB71134.1; -
 DR EMBL: AL031018; CAA19808.1; -
 DR EMBL: AL161558; CAB79248.1; -
 DR PIR: T05124; T05124.
 DR HSSP: P05020; 1879.
 DR InterPro: IPR006680; Amdihydro_1.
 DR InterPro: IPR004721; DHOdimr.
 DR InterPro: IPR002195; Dihydroorotase.
 DR Pfam: PF01979; Amdihydro_1; 1.
 DR TIGRFAMs: TIGR00856; PYRC_dimer; 1.
 DR PROSITE: PS00482; DIHYDROOROTASE_1; 1.
 DR PROSITE: PS00483; DIHYDROOROTASE_2; 1.
 KW Hydrolyase; Metal-binding; Mitochondrion; Pyrimidine biosynthesis;
 KW Transit peptide; zinc.
 FT TRANSIT 1
 FT CHAIN 1 377
 FT METAL 44 44
 FT METAL 46 46
 FT METAL 130 130
 FT METAL 168 168
 FT METAL 206 206
 FT METAL 280 280
 FT METAL 280 280
 SQ SEQUENCE 377 AA; 44937 MW; 783C22B5581DEB CRC64;
 Query Match 81.0%; Score 1467; DB 1; Length 377;
 Best Local Similarity 78.4%; Pred. No. 9e-111;
 Matches 269; Conservative 11; Mismatches 33; Indels 0; Gaps 0;
 QY 1 MEISTQPDMDHLHLDGVLKAVVSHSNHFGRAIWPNNKPTTTAAVAAREATLK 60
 DB 33 MELLTQPDMDHLHLDGVLKAVVSHSNHFGRAIWPNNKPTTTAAVAAREATLK 92
 QY 61 SLTPVDSFPLNTLYLTPTSPMEIKLARSGVYVGVCLYAGATNNSQGVDTLFGKCL 120
 DB 93 ALPSSSPPLNTLYLTPTSPMEIKLARSGVYVGVCLYAGATNNSQGVDTLFGKCL 152
 QY 121 PVLQEMVHNMLLVHGVNTPVDVMDPREKVFLETVLRPLVQVFPOLKVMMEHTTIDA 180
 DB 153 PVLQEMVHNMLLVHGVNTPVDVMDPREKVFLETVLRPLVQVFPOLKVMMEHTTIDA 212
 QY 181 VYFVSCTEGFPAAVNTQPHVLRNSLFGGLOPNAYCPLVREIRREALVSAVTS 240
 DB 213 VNFVSCTEGFPAAVNTQPHVLRNSLFGGLOPNAYCPLVREIRREALVSAVTS 272
 QY 241 KRFLGTDSAPHRBRRECCGAGIYNAPVALSVYAKVFEKENALDKLEAFSTFNGPDF 300
 DB 273 KQYFLGTDSAPHERREKSCGAGIYNBPVALSLYAKVFEBAALDKLEAFSTFNGPDF 332
 QY 301 YGLPRNNSKITLSTKTPKVPESFSYASGDIIPMFAGEMDMLP 343

Db 333 YGLPRNSSKITLKKSPMKVDPVNFPPGELVPMFAGETLQWOP 375

RESULT 3

Q94J38 PRELIMINARY; PRT; 359 AA.

AC Q94J38; 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Putative dihydroorotase.
GN Name=P0481B12.10;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_taxid=3947;
OX [1]

SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
Ra U.J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
Ra Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
Ra Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
Ra Hijihiata S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
Ra Ikono M., Itoh S., Itoh T., Itoh Y., Kobayashi N., Kamlya K.,
Ra Karagawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
Ra Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
Ra Nagaaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
Ra Namiki N., Negishi M., Ohta I., Ono N., Saito S., Sakai K., Shibata M.,
Ra Shimokawa T., Shomura A., Song J., Takazaki Y., Teresawa K., Teuji K.,
Ra Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
Ra Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
Ra Yano M., Jiang J., Gojobori T.,
RT "The genome sequence and structure of rice chromosome 1."
RL Nature 420:312-316(2002).
CC -1- CATALYTIC ACTIVITY. (S)-dihydroorotate + H(2)O = N-carbamoyl-L-
aspartate.
CC -1- COFACTOR. Binds 2 zinc ions per subunit (By similarity).
CC -1- PATHWAY: Pyrimidine biosynthesis, third step.
CC -1- SIMILARITY: Belongs to the DHOase family.
DR EMBL; AF003076; BAB56025.1; -.
DR HSSP; P05020; 1J79.
DR Gramene; Q94J38; -.
DR GO; GO:0004151; F:dihydroorotase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0019856; P:pyrimidine base biosynthesis; IEA.
DR InterPro; IPR006680; Dihydro_1.
DR InterPro; IPR004721; DHodimr.
DR InterPro; IPR002195; Dihydroorotase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR TIGRFAMs; TIGR00856; PYRC_dimer; 1.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKOWN_1.
DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
KW Hydrolyase; Metal-binding; Pyrimidine biosynthesis; Zinc.
SQ SEQUENCE 359 AA; 39734 MW; AD331AF3775F863 CRC64;

Query Match 80.5%; Score 1458; DB 2; Length 359;
Beet Local Similarity 78.1%; Pred. No. 4.6e-110;
Matches 268; Conservative 39; Mismatches 36; Indels 0; Gaps 0;

QY 2 ELSTQPDMDHMLRDGDVLRKAVVSHSHHFGRAIVMNLKRPITTTAAVAAYEALIKS 61
DB 12 ELITTRPDMDHMLREGVLAALVLRHSAMHFGRAIVMNLKRPVTTTBALAEYREELRA 71
QY 62 LPVDSDFNPLMTLYLTDTTSPMEIKLARESQVFGVKLYPAGATTNSODGVTLDFGKLP 121
DB 72 LPPGSNFVPLMTLYLTDTTSPMEIKLARESQVFGVKLYPAGATTNSODGVTLDFGKLP 131
QY 122 VLQGVNHNHMLPLVHGEVTNPEVDMFDEKVFIEFVLAPLVOKPQOLKVNMHVHTTIDAV 181
DB 132 VLEEMARQEMPLLVHGEVTDHVDTFDRKQVFIKILAPLVORLPKIVMEHITTMADV 191

QY 182 KEVESCTEGFVAATYTPQHLVLNRSLSFGGLQPHNYCLPVLRKREIHRALVSATVSGSK 241
DB 192 NFVESCKSGHVAATYTPQHLVLNRSLSFGGLQPHNYCLPVLRKREIHRALVSATVSGSK 251

QY 242 REFLGTSAPHDRRRKRCSCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFNGPDY 301
DB 252 QYFLGTSAPHDRRRKRCSCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFNGPDY 311

QY 302 GLPRNSSKITLKKSPMKVDPVNFPPGELVPMFAGETLQWOP 344
DB 312 GLPRNSSKITLKKSPMKVDPVNFPPGELVPMFAGETLQWOP 354

RESULT 4

Q82WF3 PRELIMINARY; PRT; 349 AA.

AC Q82WF3; 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Dihydroorotase homodimeric type (EC 3.5.2.3).
GN Name=pyrc; OrderedlocusNames=NE0727;
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OC NCBI_taxid=915;
OX [1]

SEQUENCE FROM N.A.
RX STRAIN=ATCC 19718 / IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
DOI=10.1128/JB.185.9.2759-2773.2003;
RA Chain P., Lamedin J.B., Larimer F.W., Regala W., Lao V., Land M.L.,
Ra Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayaveda-Soto L.A.,
Ra Arciero D.M., Holmes N.B., Whitaker M.M., Atp D.J.,
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
obligate chemolithoautotroph Nitrosomonas europaea."
RL J. Bacteriol. 185:2759-2773(2003).
CC EMBL; BX321858; CAD84638.1; -.

DR HSSP; P05020; 1J79.
DR GO; GO:0004151; F:dihydroorotase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0019856; P:pyrimidine base biosynthesis; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR004721; DHodimr.
DR InterPro; IPR002195; Dihydroorotase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR TIGRFAMs; TIGR00856; PYRC_dimer; 1.
DR PROSITE; PS00482; DIHYDROOROTASE_1; 1.
DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
KW Complete proteome; Hydrolyase.
SQ SEQUENCE 349 AA; 39083 MW; C2603ECB255D6703 CRC64;

Query Match 59.8%; Score 1083; DB 2; Length 349;
Beet Local Similarity 59.9%; Pred. No. 1.3e-79;
Matches 205; Conservative 54; Mismatches 79; Indels 4; Gaps 3;

QY 2 ELSTQPDMDHMLRDGDVLRKAVVSHSHHFGRAIVMNLKRPITTTAAVAAYEALIKS 61
DB 7 KLITTRDDMDHMLRDGNARSLVLRDARFARAIIMNLKRPVTTTEQAAAYRVLISA 66
QY 62 LPVD-SDFNPLMTLYLTDTTSPMEIKLARESQVFGVKLYPAGATTNSODGVTLDFGKC 119
DB 67 LPAELAGRFBPLMTLYLTDTTSPMEIKLARESQVFGVKLYPAGATTNSODGVTLDFGKC 125
QY 120 LPLVQGVNHNHMLPLVHGEVTNPEVDMFDEKVFIEFVLAPLVOKPQOLKVNMHVHTTID 179
DB 126 EATLEKREBDMLPLVHGEVVDPAVDYFDRKQIITIDVLPFLQRFGLRVFPHITTR 185
QY 180 AVKEFSECTEGFVAATYTPQHLVLNRSLSFGGLQPHNYCLPVLRKREIHRALVSATVSG 239
DB 186 AVEFVQTAAPNR-IAATITTAHMLNRLNLFQGLRPHHYCLPVLRKREIHRQALDEAATSG 244

QY 240 SKREFLGTDSAPHDRRRECCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSNGPD 299
 DB 245 HSREFLGTDSAPHDRRRECCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSNGPD 304
 QY 300 FYGLPRNNSKIKLSKTPWKVPSFSYASGDIIPMFAGEMLDW 341
 DB 305 FYGLPRNNSKIKLSKTPWKVPSFSYASGDIIPMFAGEMLDW 346

RESULT 5
 ID PYRC RALSO STANDARD; PRT; 344 AA.
 AC 08Y249;

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Dihydroorotase (EC 3.5.2.3) (DHORase)
 GN Name=pyrc; OrderedLocNames=RSC0487; ORFNames=RS05822;
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;

RA Salanoubat M., Genin S., Artiguenave F., Gonzy J., Mangenot S.,
 RA Ariat M., Billaut A., Broctier P., Camus J.C., Caticolico L.,
 RA Chaudier M., Choise N., Claudel-Renard C., Cunne S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiek T.,
 RA Siglier P., Thebault P., Whalen M., Winker P., Levy M.,
 RA Weisenbach J., Boucher C.A.,
 RT Genome sequence of the plant pathogen Ralstonia solanacearum";
 RT Nature 415:497-502(2002).
 CC -1 CARBAMYL ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-
 CC aspartate.
 CC -1 COFACTOR: Binds 2 zinc ions per subunit (By similarity).
 CC -1 PATHWAY: Pyrimidine biosynthesis; third step.
 CC -1 SUBUNIT: Homodimer (By similarity).
 CC -1 SIMILARITY: Belongs to the DHORase family. Subfamily 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL; AL646059; CAD14015.1; -.
 DR HSP; P05020; J079.
 DR HAMAP; MF_00219; -; 1.
 DR InterPro; IPR006680; Amidohydro_1.
 DR InterPro; IPR004721; DHODimr.
 DR InterPro; IPR002195; Dihydroorotase.
 DR Pfam; PF01979; Amidohydro_1; 1.
 DR TIGRFAMs; TIGR00856; Pyrc_dimer; 1.
 DR PROSITE; PS00482; DIHYDROOROTASE_1; 1.
 DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
 DR Complete proteome; Hydroxylase; Metal-binding; Pyrimidine biosynthesis;
 KW ZINC.
 FT METAL 14 14 zinc 1 (By similarity).
 FT METAL 16 16 zinc 1 (By similarity).
 FT METAL 100 100 zinc 1 and 2 (By similarity).
 FT METAL 137 137 zinc 2 (By similarity).
 FT METAL 175 175 zinc 2 (By similarity).
 FT METAL 248 248 zinc 1 (By similarity).
 SQ SEQUENCE 344 AA; 37816 MW; 48E7FD71912CF610 CRC64;

Query Match 58.6%; Score 1061; DB 1; Length 344;
 Best Local Similarity 60.2%; Pred. No. 8e-77;
 Matches 204; Conservative 45; Mismatches 88; Indels 2; Gaps 2;

QY 3 LSTIOPDDMHILRDGVILKAVYSASHHFGRAIYMPKIPITTTAAVAAYREALIKSL 62
 DB 5 LITVRPDDMHILRDGDALADVVGDTARQFGRAIIMPILKPPVTTTAAQRAVERILAL 64
 QY 63 PVDSDPNPLMTLYLTDTTSPMEIKLARSQVGVKLYPAGATMSOGVYTDLFEKCLPV 122
 DB 65 PACTRFEPLMTLYLTDTNTTPEEVRAARASGFHSGVLYPAGATMSDAGVTDL-RRCAKT 123
 QY 123 LOEMVEHNPPLVHGVTNPEVDMFDRKVFIEYLRPLVQKFPOLKVMEHVTITDAVK 182
 DB 124 LEMQDVGMPLVHGVTNPEVDMFDRKVFIEYLRPLVQKFPOLKVMEHVTITDAVK 183
 QY 183 FVESCTEGFVAATVTPHVLNRLNSELFOGGLQPHNYCLPYLKRETHREALVSATSSGR 242
 DB 184 YVRD-AQGVGATITFAHLLVNRNMLFVGIRPHYCLPYLKRETHREALVAAATSSGHR 242
 QY 243 FFLGTDSPAHKGLRRECCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFGPDYFG 302
 DB 243 FFLGTDSPAHKGLRRECCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFGPDYFG 302
 QY 303 LPRNNSKIKLSKTPWKVPSFSYASGDIIPMFAGEMLDW 341
 DB 303 LPRNNSKIKLSKTPWKVPSFSYASGDIIPMFAGEMLDW 341

RESULT 6
 ID Q62H45 PRELIMINARY; PRT; 352 AA.
 AC Q62H45;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE Dihydroorotase, homodimeric type (EC 3.5.2.3).
 GN Name=pyrc; ORFNames=BNA2422;
 OS Burkholderia mallei ATCC 23344.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia.
 OX NCBI_TaxID=243160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 23344;
 RA Nierman W.C., Deshaizer D., Kim H.S., Tetteelin H., Nelson K.E.,
 RA Feldblyum T., Ulrich R.L., Ronning C.M., Brinker L.M., Daugherty S.C.,
 RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
 RA Gwinn M.L., Haft D.H., Khouli H., Kolonay J.F., Madupu R.,
 RA Mohammed Y., Nelson W.C., Radune D., Romero C.M., Sarría S.,
 RA Selengut J., Shambin C., Sullivan S.A., White O., Yu Y., Zafar N.,
 RA Zhou L., Fraser C.M.;
 RT "Structural flexibility in the Burkholderia mallei genome";
 RT Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
 RL EMBL; CP000010; AAU49749.1; -.
 KW Hydroxylase.

QY 3 LSTIOPDDMHILRDGVILKAVYSASHHFGRAIYMPKIPITTTAAVAAYREALIKSL 62
 DB 9 LITVRPDDMHILRDGDALADVVGDTARQFGRAIIMPILKPPVTTTAAQRAVERILAL 64
 QY 63 PVDSDPNPLMTLYLTDTTSPMEIKLARSQVGVKLYPAGATMSOGVYTDLFEKCLPV 122
 DB 65 PACTRFEPLMTLYLTDTNTTPEEVRAARASGFHSGVLYPAGATMSDAGVTDL-RRCAKT 123
 QY 123 LOEMVEHNPPLVHGVTNPEVDMFDRKVFIEYLRPLVQKFPOLKVMEHVTITDAVK 182
 DB 124 LEMQDVGMPLVHGVTNPEVDMFDRKVFIEYLRPLVQKFPOLKVMEHVTITDAVK 183
 QY 183 FVESCTEGFVAATVTPHVLNRLNSELFOGGLQPHNYCLPYLKRETHREALVSATSSGR 242
 DB 184 YVRD-AQGVGATITFAHLLVNRNMLFVGIRPHYCLPYLKRETHREALVAAATSSGHR 242
 QY 243 FFLGTDSPAHKGLRRECCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFGPDYFG 302
 DB 243 FFLGTDSPAHKGLRRECCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFGPDYFG 302

Query Match 58.4%; Score 1058; DB 2; Length 352;
 Best Local Similarity 58.4%; Pred. No. 1.5e-77;
 Matches 199; Conservative 45; Mismatches 95; Indels 2; Gaps 1;
 QY 3 LSTIOPDDMHILRDGVILKAVYSASHHFGRAIYMPKIPITTTAAVAAYREALIKSL 62
 DB 9 LITVRPDDMHILRDGDALADVVGDTARQFGRAIIMPILKPPVTTTAAQRAVERILAL 64
 QY 63 PVDSDPNPLMTLYLTDTTSPMEIKLARSQVGVKLYPAGATMSOGVYTDLFEKCLPV 122
 DB 65 PACTRFEPLMTLYLTDTNTTPEEVRAARASGFHSGVLYPAGATMSDAGVTDL-RRCAKT 123
 QY 123 LOEMVEHNPPLVHGVTNPEVDMFDRKVFIEYLRPLVQKFPOLKVMEHVTITDAVK 182
 DB 124 LEMQDVGMPLVHGVTNPEVDMFDRKVFIEYLRPLVQKFPOLKVMEHVTITDAVK 183
 QY 183 FVESCTEGFVAATVTPHVLNRLNSELFOGGLQPHNYCLPYLKRETHREALVSATSSGR 242
 DB 184 YVRD-AQGVGATITFAHLLVNRNMLFVGIRPHYCLPYLKRETHREALVAAATSSGHR 242
 QY 243 FFLGTDSPAHKGLRRECCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFGPDYFG 302
 DB 243 FFLGTDSPAHKGLRRECCGCGAGIYNAPVALSVYAKVEKENALDKLEAFTSFGPDYFG 302

```
QY 241 KPEFLGTDSAPHDRRRKSCGCGAGIYNAPVALSVYAKVPKKNALDKLEAFPSNPGDPF 300
DB 249 PPEFLGTDSAPHAKGAKKAACGACGCTALHMLLELYAEAFDQAGALDKLEGFASFPGADF 308
QY 301 YGLPRNSKIKLSKTPMKVPESSVYASGDIIPFAEMDM 341
DB 309 YGLPRSAETVTLRRTWELPREIDAGAPVPLRGEALGW 349

RESULT 7
ID Q63QW0 PRELIMINARY; PRT; 352 AA.
Q63QW0
ID Q63QW0;
25-OCT-2004 (TrEMBLrel. 28, Created)
25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Dihydroorotase (EC 3.5.2.3).
Name=pyrC; ORFNames=BPSL2914;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
[1]
SEQUENCE FROM N.A.
RP STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Tibbali R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cheever I.,
RA Chillingworth T., Cronin A., Crossell B., Davis P., Deshaizer D.,
RA Felling W.T., Fraser A., Hance Z., Hauser H., Holtroyd S., Jørgensen K.,
RA Keith K.B., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabbittowitch E., Rutherford K., Sanders M., Simmonds M.,
RA Songerivai S., Stevens K., Tumapa S., Vessetrachayest M.,
RA Whitehead S., Yeates C., Barrett B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH36924.1; -.
KW Hydrolyase.
SQ SEQUENCE 352 AA; 38177 MW; 88B2B9B04EB1F241 CRC64;

Query Match 58.4%; Score 1058; DB 2; Length 352;
Best Local Similarity 58.4%; Pred. No. 1.5e-77;
Matches 199; Conservative 45; Mismatches 95; Indels 2; Gaps 1;

QY 3 LSTTPDDMHLHLDGDLKAVVSHSAHFGRALVMPNLKPPITTTAAAVYREALIKSL 62
DB 9 LTLARPDDMHLHVDGAMLAVALPHTARQFGRALIMPLKPPVTTTAAQAVRERRILAAV 68
QY 63 PVDSDFRPLMTLYLTDTTSPMEIKLARESOVVGKLYPAGATTNSODGVTDLFGKCLPV 122
DB 69 PAGMTFEPPLMTLYLTDTNPADRIARRRESGVHGKLYPAATTNSDAGVTDLGGCAKT 128
QY 123 LOEWEHNMPLLVHGEVTNPEVDMFDRKVFIEFVLAPLVOKFPOLKVMHEHTTIDAVK 182
DB 129 LEMOEVGMPLLVHGEVTNPEVDMFDRKVFIDRKMVILDRVMEELRGLGKLVFHHITTKAAD 188
QY 183 FVE--SCTEGFVAATVTPQHLVLNRNSLFGGGLQPHNYCLFVLKREIHRALVSAVTSGS 240
DB 189 YVRDADAASGRIGATITTAHLLVYRNAMFPGGIRPHYVCLVPLKRETHIRILVRAATSGN 248
QY 241 KRFLGTDSAPHDRRRKSCGCGAGIYNAPVALSVYAKVPKKNALDKLEAFPSNPGDPF 300
DB 249 PPEFLGTDSAPHAKGAKKAACGACGCTALHMLLELYAEAFDQAGALDKLEGFASFPGADF 308
QY 301 YGLPRNSKIKLSKTPMKVPESSVYASGDIIPFAEMDM 341
DB 309 YGLPRSAETVTLRRTWELPREIDAGAPVPLRGEALGW 349
```

```
RESULT 8
ID PYRC_SVNY3 STANDARD; PRT; 342 AA.
PYRC_SVNY3
ID PYRC_SVNY3
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
05-JUL-2004 (Rel. 44, Last annotation update)
DE Dihydroorotase (EC 3.5.2.3) (DHOase).
Name=pyrC; Ordered locusName=blr0406;
OS Synecocystis sp. (Strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
[1]
SEQUENCE FROM N.A.
MEDLINE=97061201; PubMed=8905231;
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Saenamoto S., Kimura T.,
RA Hosouchi T., Matsuo A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Matanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. 11. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-
CC aspartate.
CC -1- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC -1- PATHWAY: Pyrimidine biosynthesis; third step.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SIMILARITY: Belongs to the DHOase family. Subfamily 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC
DR EMBL; D90915; BAA18539.1; -.
DR PIR; S76410; S76410.
DR HSSP; P05020; J379.
DR HAMAP; MF_00219; -.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR004721; DHOdmr.
DR InterPro; IPR002195; Dihydroorotase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR TIGRFAMs; TIGR00856; pyrC_dimer; 1.
DR PROSITE; PS00482; DIHYDROOROTASE_1; 1.
DR PROSITE; PS00483; DIHYDROOROTASE_2; 1.
DR Complete proteome; Hydrolyase; Metal-binding; Pyrimidine biosynthesis;
KW Zinc.
FT METAL 13 13 Zinc 1 (By similarity).
FT METAL 15 15 Zinc 1 (By similarity).
FT METAL 99 99 Zinc 1 and 2 (By similarity).
FT METAL 136 136 Zinc 2 (By similarity).
FT METAL 174 174 Zinc 2 (By similarity).
FT METAL 246 246 Zinc 1 (By similarity).
SQ SEQUENCE 342 AA; 38127 MW; E8A572814296B121 CRC64;

Query Match 58.3%; Score 1056.5; DB 1; Length 342;
Best Local Similarity 58.8%; Pred. No. 1.9e-77;
Matches 200; Conservative 52; Mismatches 85; Indels 3; Gaps 2;

QY 2 ELSTTPDDMHLHLDGDLKAVVSHSAHFGRALVMPNLKPPITTTAAAVYREALIKSL 61
DB 3 KLTLTRDDMHLHLDGDLKAVVSHSAHFGRALVMPNLKPPVRSVADAAYRERILAA 62
QY 62 LPVDSDFRPLMTLYLTDTTSPMEIKLARESOVVGKLYPAGATTNSODGVTDLFGKCLPV 121
DB 63 IPAGGGEPEPLMTLYLTNTNDEEIIAAKASQFVAAYXYPAGATTNSDFGVTIDH-RCD 121
QY 122 VLQEMVHNMPLLVHGEVTNPEVDMFDRKVFIEFVLAPLVOKFPOLKVMHEHTTIDAV 181
```

Db 122 VLAMEQVDP.LLHGEVTDSDNIDFIREVFIETKILIFIREFPEPLRVLEHITTSADV 181
 Qy 182 KEVSSCTEGFVAATVTOHVLNRNSLFGGGLQPHNYCPLVREIHRALVSATVSGSK 241
 Db 182 QFVLSANN--IAATITPQHLFGRNALFKGIGLPHFYCLPLIKREHRLSLHAATSGNP 239
 Qy 242 RFLGTDSPAPHDRRKESCGCAGIYAPVALSVKAVKEKNAIDKLEAFSTFNGDPFY 301
 Db 240 KFFLGTDSPAPHDNRKESLGGCAGCYSAHAMELIEAFESVNSLDKLEAFASFYGPDPY 299
 Qy 302 GLPRNSKITLSTKTPKVPESFSYASGDIIPMFAGEMLDW 341
 Db 300 QLEPNTAQITLTMKPNRIPALFPPESGLVPLRAGEBITW 339

RESULT 9

Q7N012 PRELIMINARY; PRT; 361 AA.
 AC 07N012;
 DT 01-MAR-2004 (T-EMBLrel. 26, Created)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
 DE Dihydroorotase (EC 3.5.2.3).
 GN Name=pyrc; OrderedLocusNames=CV4330;
 OS Chromobacterium violaceum.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Chromobacterium.
 NCBI_TaxID=536;
 (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 12472 / DSM 30191;
 RX MEDLINE=2282880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
 RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Gultarases C.T.,
 RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
 RA Alves-Gomes J.A., Andrade E.M., Araxipe J., de Araujo M.F.F.,
 RA Astolfi-Filho S., Azevedo V., Baptista A.J., Batata L.A.M.,
 RA Batista J.S., Beijo A., van den Berg C., Bogo M., Bonatto S.,
 RA Bordignon J., Brito M.M., Brito C.A., Brocchi M., Butty H.A.,
 RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
 RA Carvalho C.M.B., Cascado J.C.M., Cavada B.S., Chelate L.M.O.,
 RA Czeizynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
 RA Fantiuch F., Farias I.P., Felipe M.S.S., Ferraz L.P., Ferro J.A.,
 RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
 RA Gazimelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
 RA Gracispaglia D., Grisard E.C., Hanna E.S., Jardim S.N., Laurino J.,
 RA Leol L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
 RA Madeira H.M.F., Manlio G.P., Marinho A.Q., Martins W.S.,
 RA di Mauro S.M.Z., de Medeiros S.R.B., Meisner R.V., Moreira M.A.M.,
 RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
 RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
 RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
 RA Ramalho-Neto C.B., Reis A.M.M., Rigo L.U., Rondinelli E.,
 RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seunarez H.N.,
 RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
 RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
 RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
 RA Vettore A., Wasseem R., Zaha A., Simpson A.J.G.;
 RT "The complete genome sequence of Chromobacterium violaceum reveals
 RT remarkable and exploitable bacterial adaptability.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
 CC -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-
 CC aspartate.
 CC -1- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
 CC -1- PATHWAY: Pyrimidine biosynthesis; third step.
 CC EMBL: A5016925; AA061989.1; -.
 DR HSSP: P05020; 1J79.
 DR GO: GO:0004151; F:dihydroorotase activity; IEA.
 DR GO: GO:0016787; F:hydrolase activity; IEA.
 DR GO: GO:001856; F:pyrimidine base biosynthesis; IEA.
 DR GO: GO:006221; P:pyrimidine nucleotide biosynthesis; IEA.
 DR InterPro: IPR006680; Amidohydro_1.

DR InterPro: IPR004721; DHODImr.
 DR InterPro: IPR002195; Dihydroorotase.
 DR Pfam: P01979; Amidohydro_1; 1.
 DR TIGRFAMs: TIGR00856; pyrc_dimer; 1.
 DR PROSITE: PS00483; DIHYDROOROTASE 2; 1.
 KW Complete proteome; Hydrolase; Metal-binding; Pyrimidine biosynthesis;
 KW Zinc.
 SQ SEQUENCE 361 AA; 38985 MW; 6862FD360883FEF3 CRC64;

Query Match 58.1%; Score 1053; DB 2; Length 361;
 Best Local Similarity 59.0%; Pred. No. 3.8e-77;
 Matches 200; Conservative 46; Mismatches 91; Indels 2; Gaps 2;

Qy 3 LSTIQDDMHLRLRGDVLKAVSHAHFGRATVMPKLPITTTAAVAAYREALKSL 62
 Db 21 LTLIRPDDMHLRLRDALAAVLPTSCOMGRATVMPKLPVTVAATAAAYRRRIIAAR 80
 Qy 63 PVDSDNPMLTYLTDTSPEIKLARESOVVFVKLYPAGATTNSODGYTDLFGKCLPV 122
 Db 81 PAGSAPEPLMTLYLTDKTSADEIRKAKCGFVHGKLYPAGATTNSDGYTDL-ANMPA 139
 Qy 123 LQEMVHNMPILVHGEVTPNPEVMDPREKVFIEVLAELVQKPOLKVMEHTTTDAVK 182
 Db 140 LEAMAEAGMPLVHGEVTDADIDVDFDEAVFIEVMKPLAKLPOLKVFEHITTTREAAE 199
 Qy 183 FVESCTEGFVAATVTPQHLVNRNSLFGGGLQPHNYCPLVKEIHRALVSATVSGSKR 242
 Db 200 FVAAAPAN-VAAITTAHLILNNRNALFVGGRPHHYCLPVLKELHQAALVKAATSGSAK 258
 Qy 243 FFLGTDSPAPHDRRKESCGCAGIYAPVALSVKAVKEKNAIDKLEAFSTFNGDPFY 302
 Db 259 FFLGTDSPAPHDNRKESLGGCAGCYSAHAMELIEAFESVNSLDKLEAFASINCPAFYG 318
 Qy 303 LPRNSKITLSTKTPKVPESFSYASGDIIPMFAGEMLDW 341
 Db 319 LPAGDRIEIVKESMTVPALFPPESGLVPLRAGEBITW 357

RESULT 10

Q9KJ24 STANDARD; PRT; 342 AA.
 AC 09KJ24;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 25-OCT-2004 (Rel. 45, Last annotation update)
 DT Dihydroorotase (EC 3.5.2.3) (DHODase).
 GN Name=pyrc; OrderedLocusNames=VCA0925;
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 NCBI_TaxID=666;
 (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
 RA Heidelberg J.F., Bisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
 RA Emolaeva M.D., Vamathevan J.J., Base S., Qin H., Dregoi I.,
 RA Sellers P., McDonald L.A., Utecherback T.R., Fleischmann R.D.,
 RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
 RA Mekalanos J.J., Venter J.C., Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 CC -1- CATALYTIC ACTIVITY: (S)-dihydroorotate + H(2)O = N-carbamoyl-L-
 CC aspartate.
 CC -1- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
 CC -1- PATHWAY: Pyrimidine biosynthesis; third step.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SIMILARITY: Belongs to the DHODase family. Subfamily 1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.emb-nb.ch/announce/>
CC or send an email to license@emb-nb.ch).

CC -----
CC EMBL/ AEO04420; AAF96822.1; ALT_INIT.
CC HSSP; P05020; 1079.
CC TIGR; VCA0925; -.
CC HAMAP; MF 00219; -; 1.
CC InterPro; IPR006680; Amidohydro_1.
CC InterPro; IPR004721; DHODImr.
CC InterPro; IPR002195; Dihydroorotase.
CC Pfam; PF01979; Amidohydro_1; 1.
CC TIGRfam; TIGR00856; pyrC_dimer; 1.
CC DR PROSITE; P500482; DIHYDROOROTASE_1; 1.
CC DR PROSITE; P500483; DIHYDROOROTASE_2; 1.
CC Complete proteome; Dihydroorotase; Pyrimidine biosynthesis;
KW Zinc.
KW Zinc 1 (By similarity).
FT METAL 13 13 Zinc 1 (By similarity).
FT METAL 15 15 Zinc 1 (By similarity).
FT METAL 98 98 Zinc 1 and 2 (By similarity).
FT METAL 135 135 Zinc 2 (By similarity).
FT METAL 173 173 Zinc 2 (By similarity).
FT METAL 246 246 Zinc 1 (By similarity).
SQ SEQUENCE 342 AA; 37461 MW; 9FAD2DE6E1818DA CRC64;

Query Match 57.4%; Score 1039.5; DB 1; Length 342;
Best Local Similarity 59.0%; Pred. No. 4.5e-76;
Matches 200; Conservative 45; Mismatches 91; Indels 3; Gaps 3;

QY 3 LSTTPDDMHLHRLDGDVLRKAVVSHSAHFGRAIWMENLKRPITTTAAVAAYREALIKSL 62
DB 4 LITRRDDMHLHRLDGDVLRKAVVSHSAHFGRAIWMENLKRPITTTAAVAAYREALIKSL 63
QY 63 PVSDENPLMTLYLTDTTSPMEIKLARESOVGVKLYPAGATNNSODGVTDLFGKCLPV 122
DB 64 P-QAHFEBLMALYLTDNTSPSEIRKAKSGKVAAKLYPAGATNNSOGVTSI-KNITVP 121
QY 123 LQEMVENMPLVHGEVTNPEVDMFDRKVFIEFTVLRPLVOKFPOLKVMHEVTTIDAVK 182
DB 122 LQAMQEGMQLLHGEVTAHDIDIFREKVFLETVLRPLVOKFPOLKVMHEVTTIDAVK 181
QY 183 FVESCTEGFAVATVTPQHLVLNNSLFGGLQPHNYCLPVLRKEIHREALVSAVTSQSKR 242
DB 182 FVQAGAGN-VAGITTAHHLHFNHMLVGGIRPHFYCLPLIKRATHQALVAATSGSKK 240
QY 243 FFLGTDSAPHRRRKESCCGAGIYNAPVALSVYAKYFEKEMALDKLEAFSPNGPDPYG 302
DB 241 FFLGTDSAPHAQKREKESACGAGSYTAHAALIELVAEVEKEKLENTLAFSPNGPDPYG 300
QY 303 LPRNNSKIKLSKTPKWPESFSYASGDIIMPAGEMLDW 341
DB 301 LPRNQTITLTKQAMPVAESMPFGSDIVVPIRAGEANLW 339

RESULT 11

Q6LP17 PRELIMINARY; PRT; 342 AA.
AC Q6LP17;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative dihydroorotase.
GN Name=C4330; OrderedlocusNames=PBRA2405.
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_Taxid=74109;
RN
RP SEQUENCE FROM N.A.
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,

RA Cestaro A., Malacrida G., Simonati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations."
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR378670; CAG20789.1; -.
DR GO; GO:0004151; P.dihydroorotase activity; IEA.
DR GO; GO:0016787; P.dihydroorotase activity; IEA.
DR GO; GO:0019856; P.pyrimidine base biosynthesis; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR004721; DHODImr.
DR InterPro; IPR002195; Dihydroorotase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR TIGRfam; TIGR00856; pyrC_dimer; 1.
DR PROSITE; P500482; DIHYDROOROTASE_1; 1.
DR PROSITE; P500483; DIHYDROOROTASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 342 AA; 37603 MW; AFE28AE6FEBDA022 CRC64;

Query Match 55.9%; Score 1012.5; DB 2; Length 342;
Best Local Similarity 56.0%; Pred. No. 7e-74;
Matches 190; Conservative 57; Mismatches 89; Indels 3; Gaps 3;

QY 3 LSTTPDDMHLHRLDGDVLRKAVVSHSAHFGRAIWMENLKRPITTTAAVAAYREALIKSL 62
DB 4 LITRRDDMHLHRLDGDVLRKAVVSHSAHFGRAIWMENLKRPITTTAAVAAYREALIKSL 63
QY 63 PVSDENPLMTLYLTDTTSPMEIKLARESOVGVKLYPAGATNNSODGVTDLFGKCLPV 122
DB 64 P-QGNSFPLMWIVLTDNTSAEELHKAQSHVVAALKYRAGATNNSOGVTSI-DHIRPA 121
QY 123 LQEMVENMPLVHGEVTNPEVDMFDRKVFIEFTVLRPLVOKFPOLKVMHEVTTIDAVK 182
DB 122 LQAMQEGMQLLHGEVTAHDIDIFREKVFLETVLRPLVOKFPOLKVMHEVTTIDAVK 181
QY 183 FVESCTEGFAVATVTPQHLVLNNSLFGGLQPHNYCLPVLRKEIHREALVSAVTSQSKR 242
DB 182 FVYNAGN-VAGITTAHHLHFNHMLVGGIRPHFYCLPLIKNIHQDALVKAATSGNPK 240
QY 243 FFLGTDSAPHRRRKESCCGAGIYNAPVALSVYAKYFEKEMALDKLEAFSPNGPDPYG 302
DB 241 FFLGTDSAPHAQKREKESACGAGSYTAHAALIELVAEVEADALDKLEAFSPNGPDPYG 300
QY 303 LPRNNSKIKLSKTPKWPESFSYASGDIIMPAGEMLDW 341
DB 301 LPRNTDTITLTKGSWNPETMAFGGDEVVPIRAGEANLW 339

RESULT 12

Q7W4A4 PRELIMINARY; PRT; 359 AA.
AC Q7W4A4;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Dihydroorotase (EC 3.5.2.3).
GN Name=pyrC; OrderedlocusNames=BPJ3762;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_Taxid=519;
RN
RP SEQUENCE FROM N.A.
RA STRAIN=12822 / ATCC BAA-587;
RC MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarrega A.-M., Basham D., Baeson N., Chevreton I.,
RA Achman M., Atkin R., Baker S., Basham D., Baeson N., Chevreton I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Fellwell T., Goble A., Hamlin M., Hauser H., Holtroyd S., Jørgen K.,
RA Leather S., Moule S., Norbertzak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,

RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*."
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640434; CAE39045.1; -
 DR HSSP: P05020; 1J79.
 DR GO: GO:0004151; F:dihydrocrotonase activity; IEA.
 DR GO: GO:0016787; F:hydrolase activity; IEA.
 DR GO: GO:0019856; P:pyrimidine base biosynthesis; IEA.
 DR InterPro: IPR006680; Dihydrocrotonase_1.
 DR InterPro: IPR004721; Rhodimr.
 DR InterPro: IPR002195; Dihydrocrotonase.
 DR Pfam: PF01979; Amidohydro_1; 1.
 DR TIGRFAMs: TIGR00856; pyrc_dimer; 1.
 DR PROSITE: PS00482; DIHYDROCRONASE_1; 1.
 DR PROSITE: PS00483; DIHYDROCRONASE_2; 1.
 KW Complete proteome; Hydrolyase.
 SQ SEQUENCE 359 AA; 38969 MW; A75138BC80C7DE97 CRC64;

Query Match 55.9%; Score 1012.5; DB 2; Length 359;
 Best Local Similarity 57.3%; Pred. No. 7.4e-74;
 Matches 197; Conservative 46; Mismatches 96; Indels 5; Gaps 3;

QY 2 ELSTQPDMDHLHLDGQVUKAVVSHSAHFGRAIWMNPKPITTTAAVAYR---EAI 58
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB ELITTRPDDMHLDGSALEAVLDIARQAFARAIWPNRPPVTTTEQALAYARIEAA 72
 QY 59 LKSLPVD-SDFNPMTLYLTDTTSPMEIKLARESQVFGVLYPAGATTNSQDGVTLDFG 117
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB LKAGGDTAAFTPLMTLYLTDTNTPAEIVRAHESGOVAAYKLYPAGATTNSDAGVTLDFG 132
 QY 118 KCLPVLQEMVHNHNPMLVHGEVNTPEVDMPREKVFIEVLRPLVQKFPOLKVMMEHVT 177
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 133 KCGAALALERCQMPLVHGEVTDPAIDVDFREAVFIERWQPLRRAYPGLKVFHEHTT 192
 QY 178 IDAVKFEVSCTEGFAATVTPQHLVLNRNSLFOGLOPHNYCLPVLKREIHEALVSAYT 237
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 193 REGAHVYVD-AEGTPAATITTPQHLVYRNALFTGCVAPHYCPLIKREVHROALVEAAT 251
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 QY 238 SGRKRFPLGTD SAPDRRREKSCGAGIYNAPVALSVAKVPEKENALDKLEAFSPNG 297
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 252 SGRPRFLGTD SAPHAGLKEHACGAGCTALHAMELYTATADAVGRDLRLGSEFAPFG 311
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 QY 298 PDFTGLPRNNKIKLSTKTPMKVPESPSYASGDIIPMAGEMLDW 341
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 312 PDFYGLPRNTGTTLTRREAYEIPAEVAFGDTTLVPLAGSESLGW 355
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13

Q7WFR2 PRELIMINARY; PRT; 359 AA.
 AC Q7WFR2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Dihydrocrotonase (EC 3.5.2.3).
 GN Name=pyrc; OrderedlocusNames=BB4208;
 OS *Bordetella bronchiseptica* (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 NC NCB1_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-588;
 RA MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
 RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,
 RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,

RA Rabbintowech E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*."
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640449; CAE34572.1; -
 DR HSSP: P05020; 1J79.
 DR GO: GO:0004151; F:dihydrocrotonase activity; IEA.
 DR GO: GO:0016787; F:hydrolase activity; IEA.
 DR GO: GO:0019856; P:pyrimidine base biosynthesis; IEA.
 DR InterPro: IPR006680; Dihydrocrotonase_1.
 DR InterPro: IPR004721; Rhodimr.
 DR InterPro: IPR002195; Dihydrocrotonase.
 DR Pfam: PF01979; Amidohydro_1; 1.
 DR TIGRFAMs: TIGR00856; pyrc_dimer; 1.
 DR PROSITE: PS00482; DIHYDROCRONASE_1; 1.
 DR PROSITE: PS00483; DIHYDROCRONASE_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 359 AA; 38969 MW; A75138BC80C7DE97 CRC64;

Query Match 55.9%; Score 1012.5; DB 2; Length 359;
 Best Local Similarity 57.3%; Pred. No. 7.4e-74;
 Matches 197; Conservative 46; Mismatches 96; Indels 5; Gaps 3;

QY 2 ELSTQPDMDHLHLDGQVUKAVVSHSAHFGRAIWMNPKPITTTAAVAYR---EAI 58
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB ELITTRPDDMHLDGSALEAVLDIARQAFARAIWPNRPPVTTTEQALAYARIEAA 72
 QY 59 LKSLPVD-SDFNPMTLYLTDTTSPMEIKLARESQVFGVLYPAGATTNSQDGVTLDFG 117
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB LKAGGDTAAFTPLMTLYLTDTNTPAEIVRAHESGOVAAYKLYPAGATTNSDAGVTLDFG 132
 QY 118 KCLPVLQEMVHNHNPMLVHGEVNTPEVDMPREKVFIEVLRPLVQKFPOLKVMMEHVT 177
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 133 KCGAALALERCQMPLVHGEVTDPAIDVDFREAVFIERWQPLRRAYPGLKVFHEHTT 192
 QY 178 IDAVKFEVSCTEGFAATVTPQHLVLNRNSLFOGLOPHNYCLPVLKREIHEALVSAYT 237
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 193 REGAHVYVD-AEGTPAATITTPQHLVYRNALFTGCVAPHYCPLIKREVHROALVEAAT 251
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 QY 238 SGRKRFPLGTD SAPDRRREKSCGAGIYNAPVALSVAKVPEKENALDKLEAFSPNG 297
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 252 SGRPRFLGTD SAPHAGLKEHACGAGCTALHAMELYTATADAVGRDLRLGSEFAPFG 311
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 QY 298 PDFTGLPRNNKIKLSTKTPMKVPESPSYASGDIIPMAGEMLDW 341
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 DB 312 PDFYGLPRNTGTTLTRREAYEIPAEVAFGDTTLVPLAGSESLGW 355
 ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14

Q7WUP3 PRELIMINARY; PRT; 359 AA.
 AC Q7WUP3;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Dihydrocrotonase (EC 3.5.2.3).
 GN Name=pyrc; OrderedlocusNames=BP3035;
 OS *Bordetella pertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 NC NCB1_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RA MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
 RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsels K.,

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 00:25:15 ; Search time 767 Seconds
(without alignments)
9809.637 Million cell updates/sec

Title: US-10-070-277-1

Perfect score: 1271
Sequence: 1 ttgcaaaatgagctctca.....tgaagattcactgataaaaa 1271

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1271	100.0	1271	4	AAf61217
2	1271	100.0	1271	4	AAf61262
3	646	50.8	1743	6	AAf61459
4	281	22.1	1053	8	AAf61459
5	281	22.1	1071	10	AAf61459
6	277.2	21.6	1047	8	AAf61459
7	272.8	21.5	1053	10	AAf61459
8	272.8	21.5	110000	10	AAf61459
9	272.8	21.5	182624	10	AAf61459
10	267.4	21.0	1032	8	AAf61459
11	267.4	21.0	1041	9	AAf61459
12	265.2	20.9	1056	8	AAf61459
13	264.4	20.8	1092	8	AAf61459
14	264.2	20.8	110000	6	AAf61459
15	245.6	19.3	1047	8	AAf61459
16	243	19.1	1068	12	AAf61459
17	242.6	19.1	1197	11	AAf61459
18	242.6	19.1	1262	11	AAf61459
19	241	19.0	1047	8	AAf61459
20	238.2	18.7	1062	8	AAf61459

21	238.2	18.7	1062	8	ACA39060	ACA39060 Prokaryot
22	238.2	18.7	92407	4	AAf28549	AAf28549 Genomic f
23	236.2	18.6	1044	8	ACA43541	ACA43541 Prokaryot
24	236	18.6	1032	10	ACA41184	ACA41184 Prokaryot
25	236	18.6	1032	10	ABE239013	ABE239013 N. gonorr
26	236	18.6	1119	3	AAf61520	AAf61520 N. mening
27	232.8	18.3	30078	3	AAf61520	AAf61520 N. mening
28	232.8	18.3	110000	3	AAf61490_06	AAf61490_06
29	232.8	18.3	110000	3	AAf61490_07	AAf61490_07
30	232.8	18.3	34980	3	AAf21608	AAf21608 Natsoria
31	231.2	18.2	1035	8	ACA41757	ACA41757 Prokaryot
32	221.6	17.4	1044	8	ACA35674	ACA35674 Prokaryot
33	221.6	17.4	1059	11	ACA45823	ACA45823 K16b1e11
34	219.2	17.2	1041	8	ACA45823	ACA45823 Prokaryot
35	218.4	17.2	1047	8	ACA45823	ACA45823 Prokaryot
36	218.4	17.2	1052	8	ACA48698	ACA48698 Prokaryot
37	183.8	14.5	642	13	ACN46418	ACN46418 Cotton pr
38	177	13.9	381	12	ADP95135	ADP95135 Cotton ex
39	149	11.7	574	8	ACA41496	ACA41496 Cotton pr
40	138.8	10.9	574	13	ACN48140	ACN48140 Cotton pr
41	129.8	10.2	1266	11	ABD02404	ABD02404 Pseudomon
42	114.4	9.0	1008	8	ACA10213	ACA10213 Prokaryot
43	114.4	9.0	393	10	ACF66304	ACF66304 Prokaryot
44	105	8.3	407	8	ACA44862	ACA44862 Prokaryot
45	79.6	6.3	261	11	ABD02513	ABD02513 Pseudomon

ALIGNMENTS

RESULT 1	AAf61217	standard, cDNA, 1271 BP.
ID	AAf61217	
AC	AAf61217	
XX		
DT	25-MAY-2001	(first entry)
XX		
DE	Potato dihydro-oxotase cDNA.	
XX		
KW	Potato; dihydro-oxotase; plant; polysaccharide content; orotic acid;	
XX	uridine; starch synthesis; ss.	
OS	Solanum tuberosum.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	9..1049
FT		/*tag= a
FT		/product= "dihydro-oxotase"
XX		
XX	W0200114569-A2.	
PN		
XX		
PD	01-MAR-2001.	
XX		
PR	12-AUG-2000; 2000WO-EP007884.	
XX		
PR	20-AUG-1999; 99DE-01039688.	
XX		
PA	(BADI) BASF AG.	
XX		
PI	Ehhardt T, Stltt Nigel M, Geigenberger PL, Loeff I, Zrenner R;	
FI	Schroeder M;	
XX		
DR	WPI; 2001-202938/20.	
DR	P-PSDB; AAB70773.	
XX		
PT	Increasing the polysaccharide, especially starch, content of plants, by	
FT	transforming with a sequence encoding dihydro-oxotase.	
XX		
PS	Claim 3; Page 28-30; 32p; German.	
XX		
CC	This invention describes a novel use of a DNA sequence (I) encoding	
CC	dihydro-oxotase (II) to prepare plants which have increased	

PF 20-DEC-2001; 2001WO-EP015093.
 XX
 XX 22-DEC-2000; 2000EP-00870319.
 PR 26-FEB-2001; 2001US-0271656P.
 XX
 PA (CROP-) CROPPESIGN NV.
 XX
 XX Kanhonou RA, Serrano Salom R, Ros Palau R;
 FI WPI; 2002-583567/62.
 DR P-PSDB; AA015493.
 XX
 PT Novel sugar beet genes involved in stress response, useful for rendering
 PT crops resistant to stress situations like osmotic stress caused by salt,
 PT drought, cold or frost thus increasing crop yield.
 XX
 XX Claim 4; Fig 6; 95pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of sugar beet
 CC (Beta vulgaris) genes that are involved in osmotic and oxidative stress
 CC tolerance in a plant. The Beta vulgaris genes of the invention encode the
 CC proteins: casein kinase alpha catalytic subunit; dihydroorotase;
 CC translation initiation factor 1A; putative protein; and an unknown
 CC protein. The Beta vulgaris genes of the invention are useful for
 CC enhancing the osmotic and oxidative stress tolerance of a plant. The Beta
 CC vulgaris genes are also useful for stimulating plant growth. The casein
 CC kinase alpha catalytic subunit gene is useful for controlling the process
 CC of flowering in a plant. The present DNA sequence represents the Beta
 CC vulgaris dihydroorotase gene
 XX
 XX Sequence 1743 BP; 526 A; 372 C; 344 G; 501 T; 0 U; 0 Other;

Query Match 50.8%; Score 646; DB 6; Length 1743;
 Best Local Similarity 76.4%; Pred. No. 1e-176;
 Matches 793; Conservative 0; Mismatches 245; Indels 0; Gaps 0;

QY 3 GCAAAATGAGAGCTCTCAATACACAACTGATGATGGCATCTTCACTCCGATGAT 62
 DB 193 GTAAAGATGAGAGCTCTCAATACAGCCCTGATGATGATGATCTTCACTCCGATGAG 252
 QY 63 GATGTTCTTAAGGAGCTGCTCTCAATACAGCTGATGATGATGATGATGATGATGAT 122
 DB 253 GATCTCTCTGCTGCTGCTGCTCTCAATACAGCTGATGATGATGATGATGATGATGAT 312
 QY 123 CCAATTTGAAGCTCTCTCAATACAGCTGATGATGATGATGATGATGATGATGATGAT 182
 DB 313 CCGAATCTTAAGGAGCTCTCTCAATACAGCTGATGATGATGATGATGATGATGATGAT 372
 QY 183 TTGAATCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 242
 DB 373 ATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
 QY 243 ACAACGAGCTCTCAATACAGCTGATGATGATGATGATGATGATGATGATGATGAT 302
 DB 433 ACAGCAGGCTCTCAATACAGCTGATGATGATGATGATGATGATGATGATGATGAT 492
 QY 303 TTGATCCCTGCTGCTGCTCAATACAGCTGATGATGATGATGATGATGATGATGAT 362
 DB 493 TTATACCTGCTGCTGCTCAATACAGCTGATGATGATGATGATGATGATGATGAT 552
 QY 363 TGTTCACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
 DB 553 TGTCTGCTGCTGCTGCTCAATACAGCTGATGATGATGATGATGATGATGATGAT 612
 QY 423 GTTACTAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 482
 DB 613 GTTACGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 672
 QY 483 AGACGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
 DB 673 AGACCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
 QY 543 GATGCTGTTAAGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 602

DB 733 GATCTCTCAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792
 QY 603 CAACATCTGTTTGAACAGAAATGCTCTTCAAGGGGGCTTCAACCGGATTAATAC 662
 DB 793 CAGACCTGTTTGAATGAAATCTCTCTTCAAGAGGTTTCAACCGGATTAATAC 852
 QY 663 TGCCTTCCAGTCTCTCAAGAGATCCACAGGGGCACTTGTGCTGATGATCAAGT 722
 DB 853 TGTCTTCCAGTCTCTCAAGAGAAATCCATACAGGCACTTGTGCTGATGATCAAGT 912
 QY 723 GGAAGTAAAGATTTTCTTGAAGTATGATGATGATGATGATGATGATGATGATGAT 782
 DB 913 GGAAGTAAAGATTTTCTTGAAGTATGATGATGATGATGATGATGATGATGATGAT 972
 QY 783 TGTCTTCCAGTCTCTCAAGAGATCCACAGGGGCACTTGTGCTGATGATCAAGT 842
 DB 973 TGTCTTCCAGTCTCTCAAGAGATCCACAGGGGCACTTGTGCTGATGATCAAGT 1032
 QY 843 GTGTTGAAAAGAAATGCACTGCAAGAGCTTGAAGATGATGATGATGATGATGAT 902
 DB 1033 GTATTTGAAGAGCTGCTGCTTGAAGATGATGATGATGATGATGATGATGATGAT 1092
 QY 903 GATTTTGAAGAGCTGCTGCTTGAAGATGATGATGATGATGATGATGATGATGAT 962
 DB 1093 GATTTTGAAGAGCTGCTGCTTGAAGATGATGATGATGATGATGATGATGATGAT 1152
 QY 963 GTACCCGATCTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 1022
 DB 1153 GTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
 QY 1023 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1040
 DB 1213 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1230

RESULT 4
 ID ACAS3620 standard; DNA, 1053 BP.
 ACAS3620;
 ACAS3620;
 19-JUN-2003 (first entry)
 DE Prokaryotic essential gene #35277.
 DE Prokaryotic essential gene #35277.
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 OS Vibrio cholerae.
 OS
 PN WO200277183-A2.
 PD 03-OCT-2002.
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0349223P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; AB049750.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

PS Claim 14; SEQ ID NO 41490; 1766pp; English

X

Claim 14; SEQ ID NO 41490; 1766bp; English.

Sequence 1053 BP; 267 A; 299 C; 257 G; 230 T; 0 U; 0 Other;

Query Match	22.1%	Score 281;	DB: 81;	Length 1053;
Best Local Similarity	56.8%	Pred: No.	9.8e-71;	
Matches 578; Conservative	0;	Mismatches 430;	Indels 9;	Gaps 3;

QY 15 CTTCTAATATACACAACCTGATGATTGGCACTCTTCAATCCCGGATAGGGAATGTTCTTAAAG 74
 Db 34 CTGAGGATTAACAGCCCGACAGCACTGGACGTTCACTTCCGGATGGCGATTACTGGCC 93
 QY 75 GCAGTGTCTCTCACAGTGCACATCATCTTGGGAGGGCAATAGTCATGCCAAATTTGAAG 134
 Db 94 GATTACTGTACGGGAGCATACGACCGGTTAATAACGGCCGGGGCTGATCATGCTTAAACAGGTT 153
 QY 135 CCTCCTATATACTAACCACTGCTGCTGTGTGAGCATACCGGAGAGCGATATTGGAATCTTTA 194
 Db 154 CCCCCGTATACCAACCACTGAGATGGCTTTAGGCTATTCGTAAGCAGCATTAATGGCCGCTCAG 213
 QY 195 CCTGTTGATGATGATTTCAAACCTCTTATGACACTTATTTGACAGATACAACAGTCCCT 254
 Db 214 CC---ACAAAGCGCACTTTGAACCTTTGAATGGCGCTTTACTCAACGATTAACACATCACT 270
 QY 255 ATGGAATATCAAATGACCAAGAGAGAGCCAGGTCGTATTTGGGATGAAGTTGATCCCTGCT 314
 Db 271 GAGGAAATTCGCAAAAGCCAAAGGGTCAGGCAAAAGTGGTGAAGCTTAATCCAGATC 330
 QY 315 GGTGCGACGACAAATTTCTCAAGATGGAGTGAATCTTTTGGGAAAGGTTTACAGATT 374
 Db 331 GGAGCGACCAACCAACTCGGATTTCAAGCGTAC---TTCAAGCCAAAACATTTAACCCAGTT 387
 QY 375 CTACAGAAGATGGTTGAGCATATATATGCTCTGCTGGTTCATGAGAGGTTACTATCTT 434
 Db 388 TTGCAAGGAGATGAGGAAGTCGGGTATGTTGCTGTTGGTGCACGATGAATGCACAGCAC 447
 QY 435 GAGGTTGACATGTTGATGAGAAAAGATATTCATTGAACCGGTTTAAGACCGTTGGGTG 494
 Db 448 GAAGTGACATTTTTCACCGTGAAGAAACCTTCTAGAACCGGTGGTTCCGCGCATCCGG 507

Qy	493	CAGAAATTTCCAACTTGAAGTCCGTANNGAGCATGACACCATTAATGCTTAAAG	554
Db	508	AATGATTTTCCCAACTGAAATTTGTGTTGAACCATCACACCGCGATGCAATACC	566
Qy	555	TTTGTGAATCTTGGACCTGAAGATTTGTGAGCAATCTGACCCCAACATCTTGT	614
Db	568	TTTGTCCAA---CAGCAGCGCATTAACCTTGGCGGCAACGATTTACCGGCAACATTTTCTG	622
Qy	615	TTGACAGGAATTTCTCTCTTCCAGGGGCTTACAAACGATTAATTACTGCTTCAATC	674
Db	625	TTTAAACGTAACCAATGCTGTGGTCGGTAATGCGCCAACTTCTACTGTTGCCAATC	684
Qy	675	CTCAAAAGAGAGATCCACAGGAGCACTTGTGTGACTGTTAACAATGGAAGTAAAGA	734
Db	685	CTAAACGATGCGACTCAACCAACCGCTTAATGGAGCGGCAACCTCTGGCAGAGAA	744
Qy	735	TTTTTTCTTGGACTGATAGTCTCTCATGTATAGCAAGAAAGATGTTCTTGGGA	794
Db	745	TTTCTTCTATGTAACGACTCTGCCCGCAGCAAGGCGCCAAAGAACCGCTTGTGGC	804
Qy	795	TGTGCTGTAATTTACAAATGCACCTGAGCTGTGAGTATATGCGAAGGATTTGAAAG	854
Db	805	TGGCAGGTTCTTACACAGCCCATGCAAGCTTTGAGTTGTATGCGAGTGTGGAAAA	864
Qy	855	GAATAATGCACTGCACAGCTTGAAAGCATTCATGCTTCAATGACAGATTTTATGG	914
Db	865	GAAGGTAGCTAATAAATCTGGAAGGTTCCGTACTTTAATGGCCCTGAATTTCTACGGC	924
Qy	915	CTTCTTAGGAACACTCAAGAATTAAATTGAGTAAGCGCCATGGAAGTACCGGAATCC	974
Db	925	CTACCTGCAATCAAGAGACCGTCAAGCTCAACAAGAAAGCTTGCCAGTGCAGAAAGC	984
Qy	975	TTTTCTTAATGCAACAGAGATTAATTCCATGTTTGTGGTGAATATCTGCATGG	1031
Db	985	ATGCGCTTTGGCAGCGATATCTGTGGTCCAAATCCGCGCGGTGAATAATGCAAGTGG	1041

RESULT 5
ADE99827

ID	ADB99827	standard; DNA; 1071 BP.
XX		
AC	ADB99827;	
XX		
DT	12-FEB-2004	(first entry)
XX		
DE	Bacterial polynucleotide #113.	
XX		
KW	Proteus mirabilis infection; bacterial infection; antibacterial;	
KW	immunostimulant; gene; ds.	
KW		
OS	Proteus mirabilis.	
XX		
PN	US6605709-B1.	
XX		
PD	12-AUG-2003.	
XX		
PF	05-APR-2000; 2000US-00543681.	
XX		
PR	09-APR-1999; 99US-0128706P.	
XX		
PA	(GENO-) GENOME THERAPEUTICS CORP.	
XX		
PI	Breton GL;	
XX		
DR	WP1; 2003-895291/82.	
XX		
DR	P-PSDB; ADF04000.	
XX		
PT	New Proteus mirabilis polypeptides and polynucleotides, useful as	
PT	reagents for diagnosis of bacterial disease, as components of	
PT	antibacterial vaccines, as targets for antibacterial drugs, or as	
PT	bioccontrol agents for plants.	
XX		

or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPD at [ftp.wipo.int/pub/publ/pct_sequences](http://wipo.int/pub/publ/pct_sequences)

Sequence 1047 BP; 249 A; 282 C; 237 G; 279 T; 0 U; 0 Other;

Query Match 21.8%; Score 277.2; DB 8; Length 1047;
Best Local Similarity 55.9%; Pred. No. 1.2e-69;
Matches 569; Conservative 0; Mismatches 443; Indels 6; Gaps 2;

```
QY 15 CTCTCAATCACAACCTGATGATGGCATCTTCACTCCGATGGTGAATGTTCTTAAG 74
DB 22 CTAAATAATGCGCGCCCGGATGACTGGACATTCATCTAGTGATGAAATGCTCAGT 81
QY 75 GCAAGTGTCTCTCAGTGCACATCATTGGAGGAGCAATGATCCAAATTTGAG 134
DB 82 ACCGTGTGCGCTTACCTCGAGATTTGCGCGCTATTTGTAAGCCAAATCTAGCC 141
QY 135 CCTCCATCACTACACTGCTGCTGTGATGATACCGGAGGAGATATTGAATCTTTA 194
DB 142 CAGCAATTTCAAGGTTGCGAGTGTATGCTTATCGGAGCGTATTTTGAAGCGGT 201
QY 195 CCGTGTATGATGATTTTCAACCCCTTATGACCTTTATTTGACAGATACCAAGCTT 254
DB 202 CCGCGGCGCATTAATTCACCCCGTTGATGACGTTTACGATTAATGACCTTGAATGCT 261
QY 255 ATGGAATTCAAATGACAGAGAGAGGACAGGTCGATTTTGGGGTGAAGTTGATCCTGCT 314
DB 262 AAAGATTTGACACGCGTTTGTGACAGAGCGTTTTCACGCGGCAACGTATACCGGCC 321
QY 315 GGTCCACGACAAATTTCTCAAGATGATGATGATCTTTTCCGGAAGTGTTCACAGTT 374
DB 322 AATGACACCAACCACTCCACTCAGCGTGTATCTGACAT---CCGGCAATTTACCGGTG 378
QY 375 CTACAGAAATGTTGAGCAATATATGCTCTGCTGTTTATGAGAGGTTACTAATCTT 434
DB 379 TTGAAACAAATGCAAAAGATAGGATGCCCTGCTTATTCACGGTGAAGTAAAGATGCG 438
QY 435 GAGGTGATGATGTTGATGAGAAAGATATTCATTGAAACGGTTCTTAAAGCGCTTGGT 494
DB 439 GCCGTGATCATCTTTGATGCGAAGCCCGTTTATTTAGCAAAATTTTGAAGCCATTCGC 498
QY 495 CAGAAATTTCCACAATTTGAGATGATGATGAGCAATTTACACCATTTGATGCTTTAAG 554
DB 499 CAAAGTTTCCCGAATTAATTTGCTTTGAGCATATACAGACCAAAAGATGCGGAGAT 558
QY 555 TTTGTTGATCTTGTGATGAGAGATTTTGTGACAGAACTGACCCCAACATCTTGT 614
DB 559 TATGTGCTGCGAGGCA--ATCGTTTCTTGGGGGCAACGTCACGCCCAACACTTGATG 615
QY 615 TTGAACAGGAATTTCTCTTCCAGAGGGGCTTACACACGCAATTAATTTAGCTTCAAGTC 674
DB 616 TTTTAAACGCAATTCACATCTGTGTAGCGGTATTCGCCCCCACTTGTCTGCTGCAATA 675
QY 675 CTCAAAAGAGATTCACAGGAGGACACTTGTGTAGCTGTGTAACAGTGAAGTAAGA 734
DB 676 TTGAAGCGCAGACCCATCAGCAAGCATTTGCGCGGACGCGTGCAGAGTGTTCATGATGC 735
QY 735 TTTTTCCTTGGAGTATGATGCTCTCTCATGATTAACGAAGAAAGATGTTCTTGGAGA 794
```

```
DB 736 TTCTTCCTTGGAGACCGATTCAGCTCCCATGCCAAACATGTAAGATCATCTTGGCGC 795
QY 795 TGTGCTGTATTTTAAATCATTGAGCTTGTCACTATATGCGAAGGTGTTGAAG 854
DB 796 TGTGCGGGTGTATTTCAACCCCGAGGGGATTTGCTTTAATGCTTCCGTGTGGAA 855
QY 855 GAAATGCACTGACAGAGCTTGAAGCATTCATAGCTTCAATGACACCAATTTTATGGG 914
DB 856 CTGAATGCAATTCGACATCTGGAAGCGTTTGGCGCTTAAATGCGCCAGATTTATGCG 915
QY 915 CTTCCTAGGACAACTCAAGATTAAGTTAGTAAAGCCCATGAGGATCCCAATCC 974
DB 916 TTGCGTGTAAATGATGACGTTGTGATTTGTTGCACTCCATTCCTGCAAGCAGAAAG 975
QY 975 TTTTCTTATGATCAGAGATATATTTCCCATGTTGCTGTGTAATGCTCCAGCTGGT 1032
DB 976 ATCCCATTTAGGCAATGAAATCGGTATTCCTTTCCTTGGGTCMAACGCTTAATGT 1033
```

RESULT 7
ACF68579
ID ACF68579 standard; DNA; 1053 BP.
XX
AC ACF68579;
DT 20-NOV-2003 (first entry)
DE Photobadus luminescens nucleotide sequence #7046.
XX
XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough; gene; ds.
XX
OS Photobadus luminescens.
XX
PN NO200294867-A2.
XX
XX 28-NOV-2002.
PD
XX 07-FEB-2002; 2002MO-IB003040.
PF
XX 07-FEB-2001; 2001FR-00001659.
PR
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
PI
PI Duchaud B, Taourit S, Glaser P, Frangeul L, Kunet F, Danchin A;
PI Buchrieser C;
DR WPI; 2003-148459/14.
XX
XX
PT Genomic sequence of Photobadus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
XX
PS Claim 2; SEQ ID NO 7046; 1205BP; French.
PS
XX
XX The invention relates to the isolation of genes and their encoded
CC proteins from Photobadus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of *P. luminescens*
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than *P. luminescens* and are able to alter
CC response or sensitivity to toxins and antibiotics produced by *P.*
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibiotics useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that

20 AATCACAACAACCTGATGATTTGGCAATCTTCACTCCGTGATGATGATGATTTTAAGCAGT 79
27 AATCCGCCCCCTGATGACTGGCAATCATCTTTTCGATGATGATGATTAAGTAAGTAAAGTCT 86
80 TGTCTCTCAGATGACATCATCTTTGGGAGGGCAATGATCAGCCAAATTTGAAGCTCC 139
87 CGTTCCTTACACAGTCAATTACTTTGGCCAGGCTATGCTATCTTAATCTGTGTTACGCC 146
140 TATCACTACCACTGTGCTGCTGATAGCAATCCGGAGGCCGATTTGAAATCTTTACTGT 199
147 GATTAACAGAGTGTCACTGTCGCAAGGCTTTTAAAGGCGAAGATCTGGCAGCTATTCGGA 206
200 TGAATGTGATTTCAACCTCTTATGACACTTTATTTGACAGATACAAACAGTCTTAATGA 259
207 AGAGATTAATTTCCAGCTCTTATGACTTGTATCTTAATCTGACACAGACTGAAAGCTCTCA 266
260 AATCAAACTACCAAGAGAGGCCAGTCTGATTTGGGGTGAAGTTGTAACCTCGTGTGC 319
267 GATGAAATTTGTTATTAAGAGGCATTTTACCGCTTGCAAGCTCTACCTCGCAATGC 326
320 CACGACAATTTCTCAAGATGAGTGAATGATCTTTTCGGAGTGTATTAACAGTTCTACA 379
327 CACAACCAATTCAGTCAAGCGCTCTCTGATTAATAA--GAATATTATTCACATATTAGC 383
380 AGAATGTTGAGCATTAATATGCTGCTGCTGCTGCTGATGAGAGGTTACTTAATCTTGAGT 439
384 CGTATGGAATAACTGGCGAATGCTCTTGTCTGTGTCACAGTAGAGTTACTGCTTCTCATAT 443
440 TGAATTTTGAATGAGAAAAGGATTCATTGAAAACGGTTCTAAGACCGTGTGGTCAGAA 499
444 TGAATTTTGAATGAGAAAAGCCCGTTTATTGAAACAATCATGGAACAATAGTAATCA 503
500 ATTTCACAATTTAAGTGTGATGAGAGATTTAACCACTTAATGCTGTAAAGTTGT 559
504 ATTCGAGCACTGAATGTGTGTTGGAACATTAATCAACAAAGACCGCACAAATATGT 563
560 TGAATCTGACATGAGATTTGTTGACAGCACTGTCAACCCACCAACATCTGTTTGA 619
564 GCTGGAAGGCATTA--CAATCTGGCAGCAAACTTAATCTCACAGCATCTCATGTTCA 620
620 CAGGAATCTCTTTCAGAGGGGGCTTCAACCCGATTAATTACTGCTCCAGTCTCA 679
621 TCGCAACCAATTAATGATTGTTGGGGAATTCGGCCACACTTAATCTGTGTGCTGTAA 680
680 AAGAGAGATCCACAGGGAGGCACTTGATGAGCTGTAAACAAGTGAAGTAATAAGATTTT 739
681 AGGCATATGTGATCAAGAGGCTTTTACGGGACCGCTGCGCAGTGGCTGTGATGCGTTCT 740
740 TCTTGGGACTGATATGCTCTCTCATGATAGCAAGAAAGAAAGATGTTCTTGTGATATGTC 799
741 TCTAATGATCAGACTCAGCACCAATGCAACAAGAAAGAAAGATCCTCTTGTGTGTGGC 800
800 TGTATTTTAAATGACACCTGTAAAGCTTGTCAATATATGCAAGGTGTTGAAGAGAAA 859
801 CGGCGCTCTTAAAGCAATCAGGCGCTCACTGCTTATGCAACAAGTGTGAGAGATGA 860
860 TGCATCGACAAGCTTGAAGATATCACTAGCTTCAATGACAGATTTTAAATTTGAGCTTCC 919
861 TGCATTAACCGCATTTGAAGCTTTCTGTTCATCTTAATGACCTTAATTTTACGGGCTACC 920
920 TAGGAACAACATCAAGATTAAAGTTGATGAAGACGCCATGAAGTAAACCGAATCCTTTTC 979

WP	ACF67367_00	1	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000
WP	ACF67367_03	300001	410000
WP	ACF67367_04	400001	510000
WP	ACF67367_05	500001	610000
WP	ACF67367_06	600001	710000
WP	ACF67367_07	700001	810000
WP	ACF67367_08	800001	910000
WP	ACF67367_09	900001	1010000
WP	ACF67367_10	1000001	1110000
WP	ACF67367_11	1100001	1210000
WP	ACF67367_12	1200001	1310000
WP	ACF67367_13	1300001	1410000
WP	ACF67367_14	1400001	1510000
WP	ACF67367_15	1500001	1610000
WP	ACF67367_16	1600001	1710000
WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000
WP	ACF67367_20	2000001	2110000
WP	ACF67367_21	2100001	2210000
WP	ACF67367_22	2200001	2310000
WP	ACF67367_23	2300001	2410000
WP	ACF67367_24	2400001	2510000
WP	ACF67367_25	2500001	2610000
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
P	ACF67367_56	5600001	5648894

	Matches	554;	Conservative	0;	Mismatches	442;	Indels	6;	Gaps	2;		
QY	20	AATCA	CACAA	CTGATGATTTGGCATCTTCACTCCGTATGATGATGTTCTTTAAGGCAGT	79							
Db	43973	AATCCG	CCCCC	CTGATGATCTGGCACTATCTATTTTGGTATGGTAAAGCTTAAGAACTGT	44032							
QY	80	TGCTCT	CA	CAGGCACTACCTTTGGGGGCAATATGCAATGCCAAATTTGAAGCCCTCC	139							
Db	44033	CGTCC	CTTA	CAAGCTATTACTTTGGCCGAGCTATCGTCAATGCTTAATCTCTTAACCC	44092							
QY	140	TATCA	TACCA	CTGCTGCTGCTGTATGACATACCGGAGCGATATTGAATCTTTA	199							
Db	44093	GATTA	CAGA	AGTGTGATGCGCAAGCCTTATAGAGCGAATAATCTGGACGATATCCGA	44152							
QY	200	TGAT	TAGTATTTCAACCCCTCTTAATGACATTTATTTGACAGATTAACAACAGTCCATGSA	259								
Db	44153	AGGAG	TAATTTCCACACCTCTTATATGACTTTATCTTACTGACAGACTGAAGAGCTTCA	44212								
QY	260	AATCA	AACTAGCAAGAGAGAGACGAGTCTATTTTGGGGTGAAGTTTACCCCTGCTGTGC	319								
Db	44213	GAT	AGAAATTTGGTTATTAAGAAAGCATTTTACCGGTTCAGAGCTTACCTCGCAATGC	44272								
QY	320	CACGA	CAATTTCTCAAGATGAGTGA	CTGATCTTTTCGGGAAGTGTTAACCA	379							
Db	44273	CACAA	CCAATTTCAAGTCA	CGGCGCTCTGTATTA	44329							
QY	380	AGAA	ATGTTGACATTAATATGCTCTGCTGGTTCATGAGAGGTTACTAATCCGAGGT	439								
Db	44330	CG	ATAGAAAACTGGGGATGCTCTGCTCGCCAGGTGAGTTACTGCTTCTCATAT	44389								
QY	440	TGACAT	GTTTGATAGAGAAAGGATTTCAATGAA	CGGTTCTTAAGACGTTGGTGCAGA	499							
Db	44390	TGA	TATTTTGCATCGAAGACCGGTTTATTTGAACAAGTATGAAACCACTACGTATCA	44449								
QY	500	ATTTCC	CAATTTGAAGTCTGTATGAGACATGTTACACCATTTGATGCTGTTAAGTTGT	559								
Db	44450	ATTCC	CACTGAATGTTGTTGAA	CAATTA	CTACCAAGAAAGCGCAAAATGT	44509						
QY	560	TGAT	CTTGGACAGGAATTTGTTGAGCACTGTCA	CCCCCA	CATCTTGTTTGAA	619						
Db	44510	GCT	GAAGGCA	TA	CAATCTGACGACAACTTA	CTCCACAGACATCTATGTTCA	44566					
QY	620	CAGAA	TTCTCTCTTCMAAGGGGCTTACA	CCGCATTAATTA	CTGCTTCA	GTCTCTCA	679					
Db	44567	TCCGA	ACATATGTA	GTGTGGGATTCG	GCACA	CTTA	CTGTCTGCTGTCTAA	44626				
QY	680	AA	AGAA	ATCCACAGGAGGACATGTGTAGCTGTAA	CAAGGAA	AGTAA	AGATTTT	739				
Db	44627	ACG	AAATGTCATCMAAGGCTTTA	CGGGA	GCCTG	CCAGTGGCTGTATGCTTCTT	44686					
QY	740	TCTT	GGAGCTGATAGTCTCTCTCATGATTA	CAAGAA	AGA	TGTCTTGTGATGTGC	799					
Db	44687	TCT	AGTACAGCTCAGACCA	CACTGACAA	CAAA	AAAGAA	AGATTCCTTTGTGTGC	44746				
QY	800	TG	TA	TATTAACANTGACCTGTAGCTTGTCA	GTATG	CAAGAGT	TTGAAAA	859				
Db	44747	CGG	GTCTTTTA	CCCA	CCATCAGCGCTA	CTGCTTATG	CAACAGT	TTTGA	GAAGATGAA	44806		
QY	860	TG	CA	CTGACAGAGCTTAAGCATCTAC	TAGCTTCAAT	GAAGCA	AGATTTTA	TGAGGCTTCC	919			
Db	44807	TG	CA	TTAACCGCATTTTGA	AGCTTTCTGTTCA	CTTAATG	GCCTTAATTTTA	CGGGCTACC	44866			
QY	920	TAG	AA	CAAACTCAAA	GAATTA	GTAGTAA	AGCGCATG	AGAGTAC	CCGAATCTTTTC	979		
Db	44867	TGT	CA	AA	GAAGGCTTTATTTGA	ACTTA	CTGCAAA	CTTCA	ACGCTCAT	TGAACA	ATGGA	44922
QY	980	TTA	T	ATGATCAGAGATATATTTCC	ATGTTTGTCTGTGA	AAATG	CTGAC	ATGG	1031			
Db	44927	TTG	CA	TATATATAAAAA	CTTATTTCA	ATTTTGGCTGGG	AGAAATG	CCGCGTGG	44978			

XX	ACF65379 standard; DNA; 182624 BP
XX	AC ACF65379;
XX	DT 20-NOV-2003 (first entry)
XX	DE Photorhabdus luminescens nucleotide sequence #32.
XX	KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW	detection; food; gene expression; plant; animal; microorganism; toxin;
KW	antibiotic; biopesticide; virulence factor; disease model; plague;
KW	whooping cough; gene; ds.
OS	Photorhabdus luminescens.
FN	W0200294867-A2.
PD	28-NOV-2002.
PP	07-FEB-2002; 2002MO-IB003040.
PR	07-FEB-2001; 2001FR-00001659.
PA	(INSP) INST PASTEUR.
PA	(CNRS) CNRS CENT NAT RECH SCI.
P1	Duchaud E, Taourit S, Glaeser P, Frangeul L, Kunst F, Danchin A,
DR	Buchrieser C;
PT	WPI; 2003-148459/14.
PT	Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PS	useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX	Claim 1; SEQ ID NO 32; 1205bp; French.
CC	The invention relates to the isolation of genes and their encoded
CC	proteins from Photorhabdus luminescens. The isolated sequences are
CC	sources of probes and primers for detecting the genome of P. luminescens
CC	and related species; to study polymorphisms; for gene analysis and for
CC	detection/amplification of the genes. Antibodies (Ab) raised against the
CC	polypeptides encoded by the genes are used for detection/identification
CC	of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC	carry a gene-containing vector are used to select compounds that
CC	modulate, regulate, induce or inhibit expression of the genes in plants,
CC	animals or microorganisms other than P. luminescens and are able to alter
CC	response or sensitivity to toxins and antibiotics produced by P.
CC	luminescens. Cells transformed to express the genes are useful for
CC	recombinant production of the proteins, particularly toxins and
CC	antibacterials useful as insecticides, bactericides and fungicides. The
CC	genes, proteins, vectors containing the genes and Ab are also useful
CC	therapeutically (to treat microbial infection by bacteria or fungi that
CC	are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC	biopesticides. Other uses of the genes and the proteins are as virulence
CC	factors and for identifying targets of human diseases for which P.
CC	luminescens is a model (particularly plaque and whooping cough). This
CC	sequence represents one of the isolated P. luminescens genes
XX	Sequence 182624 BP; 52367 A; 35105 C; 41128 G; 54021 T; 0 U; 3 Other;
QY	Query Match 21.5%; Score 272.8; DB 10; Length 182624;
Db	Best Local Similarity 55.7%; Pred. No. 2.7e-67; Mismatches 442; Indels 6; Gaps 2;
QY	Matches 564; Conservative 0;
Db	20 AATCACAACAAGCTGATGATTGGCATCTTCATCTCGTGATGTGATGTTCTTAAGCACT 79
77563	AATCGGCCCGCCCTATGATGATGCACATTCATTTTTTGGTAGTAGGAATCTAAAGACTGT 77504
QY	80 TGCTCTGCAGAGTGACATCATCTTTGGAGGGCAAATGTCATGCCAAATTGAAAGCTCC 139
Db	CGTTCCTCACACAGATCATTACTTTGGCCGACAGCTATCGTCATCTCCAAATCTGCTTAAGCC 77444
QY	140 TATACATAACAATGCTGCTGTGTGATGATACCGGAGGAGGCAATTAAGAATCTTTAACCCTG 199

Db	77443	GATTACGAAGTGTGACGTGACCAAGGCTTAATAGAGGCAGATATCTGGCAGCTATTTCGCCA	77384
Qy	200	TGATAGTATTTTCAMCCCTCTTATGACACTTTAATTGACAGATTCAMCAAGTCTTATGGA	259
Db	77383	AGAGAGATTAATTTCCAGCTCTTATGACTTGTATCTTACTGACACGACTGAAAGCTCTCA	77324
Qy	260	AATCAAACTAGCAGAGAGAGGACGAGTGTATTTGGGGTATAGTTGTATACCTCTGGTGC	319
Db	77323	GATAGAAATTGGTATTAAAGAAAGGCAATTTTACCGCTTGCAGCTCACTCCGCAATGC	77264
Qy	320	CACGACAAATTTCTCAAGATGAGATGACCTGATCTTTTGGGAATGTTTACAGTCTTACA	379
Db	77263	CACACCAATTAAGTCAACGGGCTCTGTGATATAA---GAATATTATCACTATAGC	77207
Qy	380	AGAAATGGTGTAGCATTAATATGCTCTGCTGCTGCTTATGAGAGGTTACTTAATCTGAGGT	439
Db	77206	CGTATGAAAAAATCTGGGAGATGCTCTGCTGCTGCACAGGAGGATTAATGCTTCTCATAT	77147
Qy	440	TGACATGTTTGTATAGAAAAGAAAGATTCATTGGAACGGTCTTAAAGCGTTGGTGCACAA	499
Db	77146	TGATATTTTTCATCAGAAAGCCCGTTTATTGAAACAAGTCATGAAACCACTACGTAACA	77087
Qy	500	ATTTCACAATTAAGAGTGTGATGAGAGCATGTTTACCACATTATGCTGTTAACTTGT	559
Db	77086	ATTCCACACACTGAAAATGTTGTGTTGAAACAATACTTACAAAGAAAGCCGCAAAATG	77027
Qy	560	TGATCTTTCGACTAGAGATTTGTTGAGCACTGTCAACCCACAACTCTGTTTGA	619
Db	77026	GCTGAAAGGCATATA---CAATCTGGGAGCAATTAATCTCACAGCATCTCACTGTTCA	76970
Qy	620	CAGAAATCTCTCTTCCAAAGGGGCTTACAAACCGCATATTATCTGCTTCCAGTCTCAA	679
Db	76969	TGCGAACCATATGCTAATGTGTGTGGATTGGGCCAACCTTAATCTGTCTGTGCTAA	76910
Qy	680	AAGAGATTCACAGGAGGCACTTGTGTCACTGTAAACAAGTGGAAAGTAAATAATTTT	739
Db	76909	ACGCAATGTGCATCAAGAGGCTTTACGGGCAACCGTGCAGTGGCTGTATGCTTCTT	76850
Qy	740	TCTTGGGACATGATAGTCTCTCATGATATGAGAGAAAGAAAGAGTCTTGTGGATGTGC	799
Db	76849	TCTAGTATACGACTCAGACACATCAGACACAAAGAAAGAAATCCCTTGTGTGTGTGC	76790
Qy	800	TGGAATTTACATGACACTGTAGCCCTGTCAATATATCGAAGGTGTTGAAAAGAAAA	859
Db	76789	CGGGTCTTTTAAAGCACATCAGCCCTACCTGCTTATGCAACAGTGTTTAGAGATGAA	76730
Qy	860	TGCACTCGACAAAGCTTGAAGACATTCATAGCTTCAATGACACAGATTTTATGGGCTTCC	919
Db	76729	TGCATTAACGCAATTTGAAAGCTTTCTGTCTCACTTAAATGGCCCTTAAATTTTACGGGCTACC	76670
Qy	920	TAGGAACAACCTCAAAAGATTAAGTTGATGATAGACGCCATGAGAGTAAACCGGAATCTTTTC	979
Db	76669	TGTCAATGAAAGCTTATTTGAATCACTACGCAAACTTTTACGTCATTTAGAACACTTGA	76610
Qy	980	TTATGATCAGAGATTAATTAATCCATGTTTGTGCTGGTAAATGCTGACTGG	1031
Db	76609	TTGCATTAATGAAAAACTTATTCATTTTGGCTGGGGAAGATGCCCGCTGG	76558

RESULT 10
ACAA20646
ID ACAA20646 standard; DNA; 1032 BP.
XX
XX AC
XX ACAA20646;
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #2303.
XX
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
drug design; gene.

OS	Acinetobacter baumannii.
XX	WO200277183-A2.
PN	
PD	03-OCT-2002.
XX	
PF	21-MAR-2002; 2002WO-US009107.
XX	
PR	21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-0362699P.
XX	
PA	(ELIT-) ELITRA PHARM INC.
XX	
P1	Mang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	
DR	WPI; 2003-029926/02.
XX	P-PSDB; ABU16776.
XX	
PT	New antisense nucleic acids, useful for identifying proteins or screening
PT	for homologous nucleic acids required for cellular proliferation to
XX	isolate candidate molecules for rational drug discovery programs.
XX	
XX	Claim 14; SEQ ID NO 8516; 176pp; English.
XX	
CC	The invention relates to an isolated nucleic acid comprising any one of
CC	the 6213 antisense sequences given in the specification where expression
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:
CC	(1) a vector comprising a promoter operably linked to the nucleic acid
CC	encoding a polypeptide whose expression is inhibited by the antisense
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC	polypeptide or its fragment whose expression is inhibited by the
CC	antisense nucleic acid; (4) an antibody capable of specifically binding
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC	proliferation or the activity of a gene in an organism required for
CC	proliferation; (7) identifying a compound that influences the activity of
CC	the gene product or that has an activity against a biological pathway
CC	required for proliferation, or that inhibits cellular proliferation; (8)
CC	identifying a gene required for cellular proliferation or the biological
CC	pathway in which the proliferation-required gene or its gene product lies
CC	or a gene on which the proliferation-required gene or its gene product lies
CC	an organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC	compound's activity; (11) a culture comprising strains in which the gene
CC	product is overexpressed or underexpressed; (12) determining the extent
CC	to which each of the strains is present in a culture or collection of
CC	strains; or (13) identifying the target of a compound that inhibits the
CC	proliferation of an organism. The antisense nucleic acids are useful for
CC	identifying proteins or screening for homologous nucleic acids required
CC	for cellular proliferation to isolate candidate molecules for rational
CC	drug discovery programs, or for screening homologous nucleic acids
CC	required for proliferation in cells other than S. aureus, S. typhimurium,
CC	K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC	prokaryotic essential genes. Note: The sequence data for this patent lies
CC	not form part of the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pat_sequences
XX	
SQ	Sequence 1032 BP; 309 A; 210 C; 205 G; 308 T; 0 U; 0 Other;
Query Match	21.0%; Score 267.4; DB 8; Length 1032;
Best Local Similarity	56.4%; Pred. No. 8,7e-67;
Matches 540; Conservative	0; Mismatches 411; Indels 6; Gaps 2;
OY	27 CAACCTGATGATTGGCATCTTCATCTCCGATGTGATAGTCTTAAGCAGTTGCTCT 86
DZ	22 CAGCGAATGATGTGCATGACACTTTACGTGATGGTTTGACATTAAGAAGTACTTCCA 81
OY	87 CACACTGCACATCACTTTGGAGGGCAATGATGCCAAATTGAAGCTCTCATCACT 146
DZ	82 GATTGGCTAACAATTCGCCCGTGGCATTTGTATGCCCTAACCTGTATCCGGCTGTAAA 141

Oy	147	ACCACTGCTGCTGCTGATGACATACCGGAGGGGAGATTTGAAATCTTTACTGTTGATG	206
Db	142	ACAGTAGAAGAAGCTTTAGACTTATCGCGAAGCATCTTGCTCATGTTTCGGAAAGGCAT	201
Oy	207	GATTTCAACCTCTTATGACACTTTATTTGACAGATACCAACAGTCTTATGAAATCAAA	266
Db	202	AATTTGACCCCTCGATGATGGCTTTATTTTAACTGACCAACACTTCACACAGATGAATCGT	261
Oy	267	CTAGCAAGAGAGACCGAGTCCGATTTTGGGGGGAAGTTGACCTCGTGGTCACAGCA	326
Db	262	AAAAATTAAAGATCAGAACATGTAATGAGATTTAACTTTATCTCTGCTGGTCGACACA	321
Oy	327	AATTCACAAGATGAGATGACTGATCTTTTCGGGAATGTTTACAGTTCTACAAGAATG	386
Db	322	AACTCCGATATAGTGTGTGATGATTT---CGTAAAGTTTATGACAGTCAATGACATTA	378
Oy	387	GTTGAGCATATATATGCTCTGCTGCTTCATGAGAGGTTACTAATCCGAGTTGACATG	446
Db	379	GAAGGACATCAAGTCCGTTATGCTTCATGAGGATTAATCAATCATCTGATGATAT	438
Oy	447	TTTGATAGAAAAAGATATTCATTGAAACGGTCTTAAGACCGTTGGGTCAGAAATTTCCA	506
Db	439	TTTGATCGGAAAAAAGCTTCCATGATGAGATTTATCACTCTATTTGAAACAGTCCG	498
Oy	507	CAATTGAAGTGGTATGAGCATGTTTACACCAATGATCTGTTAAGTTTGTGTAATCT	566
Db	499	AAACTTAAAGTGTGCTTGAAGCATCATCAACAGTATGACACACACTTGGT---TTA	555
Oy	567	TGCACGTGAAGATTTGTGCAAGCACTGTGACCCACACACACTTGTGTTGAACGAAAT	628
Db	556	GAACAAAGACCGTATATGTGCGCGCAACATCATCTCCACACATTTATTTAAACCGTAA	615
Oy	627	TCTCTCTTCCAGGGGGCTTACAAACCGCAATTAATCTGCGTCCAGTCCCAAAAGAG	688
Db	616	GATATGTTGGTGTGTGTATTTAAACCGCACTTTTACTGTTTACCAATTTTAAACGTCAA	673
Oy	687	ATCCACAGGAGGACCTTGTGTAGCTGTAAACAAGTGAAGTAAAGATTTTTTTCTTGGG	746
Db	676	ACACATCAAAACGACTTGTCTTGAAGTGGCAGAACGGTAAATCTTAATTTTTCTTAGGT	733
Oy	747	ACTGATATGTCCTCTCATGATTAAGACGAAGAAAAGTGTCTTGTGATGTGCTGTAAT	806
Db	736	ACAGACAGCGCTCCTCATGCAAAAATGCAAAAAGAGATCTGTGTGATGTGCAAGGCTGC	795
Oy	807	TACATGCACTGTAGCTTGTCAAGTATATGCCAAGGTGTTGAAAAGAAATGCACTC	866
Db	796	TATATGTCACCAAAATGCAATTTGAGCTTTATGACACACATTTGCAAGTGGTAATTA	855
Oy	867	GACAACTGGAACATTCATAGCTTCAATGACACCAATTTTATGAGGCTCTCTAGAAC	928
Db	856	GAGCGCTTGAAGAGTTTGTGATGCCATTTTGTGTGGGACTTCTACGGCTACACAGTAA	911
Oy	927	AACTCAAGATTAAGTTGAGTAAAGCCGCAATGAAAGTATCCGGAATCTTTCTTAT	983
Db	916	ACTTTTACCATTTACTTTGGTTAAAGAGATTAACCTCGTTCAGATCTTTTGATTTAT	972
RESULT 11			
ADA30128			
ID ADA30128 standard; DNA; 1041 BP.			
AC ADA30128;			
XX			
DT 20-NOV-2003 (first entry)			
XX			
DE DNA encoding Acinetobacter baumannii protein #1415.			
XX			
KM ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;			
XX vaccine; plant biocontrol agent.			
OS Acinetobacter baumannii.			
XX			

[illegible]

QY 567 TGAAGTGAAGATTTGTTGACAGACGTCACCCCAACATCTTTGTTGAACGGAAT 626
 Db 562 GAAACAGACCGTATATGTTGGGCAACAAATCACTCCCAACATTTATTTAATACCGTAT 621
 QY 627 TCTCTCTCCAAAGGGGCTTAAACCGCATTAATTTACTGCTTCCAGTCTCTCAAAAGAG 686
 Db 622 GAATATGTGTGCTGGTGTATTAACCGCATCTTACTGTTTACATTTTAAAGCCAA 681
 QY 687 ATCCACAGGAGGACCTGTGTGACGTGTAAAGTGAAGTAAAGATTTTCTTGGG 746
 Db 682 ACCATCAAAACGACTTACTTGAAGTGAAGGACAGCGGTAATCTAAATTTTCTTAGT 741
 QY 747 ACTGATATGCTCTCTCATGATAGACGAAAGAAAGATGTTCTTGATGATGCTGAT 806
 Db 742 ACAGACAGCGCCCTCATGACAAATGCAAAAGAAAGATGCTTGATGATGAGGCTGC 801
 QY 807 TACATATGACCTGTGACCTTGTGATATGCAATATGCAAGGTTTGAAGAAATGCACTC 866
 Db 802 TATATGACCAAAATGCAATGAGCTTATGACAAAGCATTTGACAAAGTGGTAATTA 861
 QY 867 GACAACTTGAAGCATTCATGATGACGACGATTTTATGAGCTTCTTGAAGAC 926
 Db 862 GAGCGCTTGAAGGTTTGTGATGACATTTTGTGACAGCTTCTGATGACGATAT 921
 QY 927 AACTGAAATTAATGATGATGACGACGATGAAAGGTTACCCGATCTTTCTTAT 983
 Db 922 ACTTACCATTAATTTGTTTAAAGAAATTAACCTGCTTCAAGATCTTTGATAT 978

RESULT 12
 ACA26372
 ID ACA26372 standard; DNA; 1056 BP.

ACA26372;
 19-JUN-2003 (first entry)
 Prokaryotic essential gene #8029.

Antisense; ds; prokaryotic essential gene; cell proliferation;
 drug design; gene.

Burkholderia mallei.

MO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002MO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(BLIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW,
 Wall D, Traxwick JD, Carr GS, Yamamoto R, Forsyth RA, Xu HH;
 WPI, 2003-029926/02.
 P-PSDB; AB022502.

New antisense nucleic acids, useful for identifying proteins or screening
 for homologous nucleic acids required for cellular proliferation to
 isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 14242; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression

of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a gene in an operon required for
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-regulated gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is one of the target
 prokaryotic essential genes. Note: The sequence data for this patent did
 not form part of the printed specification, but was obtained in
 electronic format directly from WIPO at
 ftp.wipo.int/pub/published_sequences

Sequence 1056 BP; 165 A; 369 C; 368 G; 154 T; 0 U; 0 Other;

Query Match 20.9%; Score 265.2; DB 8; Length 1056;
 Best Local Similarity 54.4%; Pred. No. 3.8e-66;
 Matches 559; Conservative 0; Mismatches 463; Indels 6; Gaps 1;

QY 14 GCTCTCAATGACACACCTGATGATGACCTTTCATCTCCGATGATGATGATTTAA 73
 Db 24 GCTGACCTGCGCCGCCGCCAAGCACTGACGTCGCGGACGCGCGATGCTCGC 83
 QY 74 GCGATGCTCTTCAAGTACATGACATCTTGGAGGCGCAATGATGACCAATTTGAA 133
 Db 84 GCGCGTCTGCGGACACCGCTCGCCAGTTGCGCGCGGATATATGCGCACTGAA 143
 QY 134 GCTCCATGATCAATCACTGCTGCTGATGATACCGGAGCGCATATTGAACTTT 193
 Db 144 GCGCGCTGACGACGCGCGGACGCGGACGCTTACCGGACGATCTGCGCGGT 203
 QY 194 ACCTGTTGATGATGATTTCAACCTCTTATGACATTTTATGACAGATCAACAGTC 253
 Db 204 GCGCGCGGATGACGTTGAGCGCTTATGATGACGCTGATCGACAGACACCGCGC 263
 QY 254 TATGAAATTAATCTGCAAGAGAGAGCGAGTCTGATTTGGGGTGAAGTTTACCTGC 313
 Db 264 CGAGAAATTCGCGCGGACGCGAAAGCGCTGCGGACGCGGCGTATCTACCGGC 323
 QY 314 TGTGCGACGACAAATTTCAAGATGAGTGAATGATCTTTTGGGAAGTGTTCACGT 373
 Db 324 GCGCGGAGACGAACTCGGACCGCGGCTGACCACTGCTGCGCAAGTGGCGAGAC 383
 QY 374 TCTACAAAGAAATGTTGACATATATGCTGCTGTTGATGAGAGGTTACTATCC 433
 Db 384 GCTCGAGGCGATGACAGAAAGTGGATCGCGCTGCTGTCGACGCGAGTGAAGGATC 443
 QY 434 TAGAGTGAATGATTTGATGAGAAAGATATCTTAAACGGTTCTAAGACGTTGGT 493
 Db 444 GTGATGACCTGTTGACCGGAGAGATGTTATCATGATCGGTGATGAGCGCGTGG 503
 QY 494 GGAAGAAATTTCAATTAATGAGTGTGATGAGACATTTTCAACCAATTGATGTTAA 553
 Db 504 CGCGCGCTGCGGCGGCTCAAGGTGTGTTGACATATTTACAGCAAGATGCGCGCA 563
 QY 554 GTTGT-----TGAATCTTCACTGAAGGATTTGTTGACGAACTGTCAACCCCAACA 607

Db 564 CTACGTCGCCGATCCGACCGCGGCTCAGGCGCGGATCGGCGACATCACCGCGCACCA 623
Qy 608 TCTTGTGTTGAGAGGAAATTCCTCTTCCAAAGGGGGTTTAAACCGCATTAATTAATGCTCT 667
Db 624 TCTGCTGTACAAACCGCAATGCATGATGTTTTTCCGGGGGATCGGTCCACATTAATGCTCT 683
Qy 668 TCCAGTCTCAAAAAGAGATCCACAGGAGGACCTTGTCTCAGCTGTAACTAAGTGAAG 727
Db 684 GCCGGTCTCAGACGCGAGACGATCGGATCGGCTGTGAGAGCGCGACGTCCGGCAA 743
Qy 728 TAAAGATTTTTCTTGGGAGCTGATGTCCTCCATGATAGACGAAGAAAGATGTC 787
Db 744 TCCGCGCTTCTCTCGGACCGACAGCGCGCGACGAAAGCGCAAGAGCGCC 803
Qy 788 TTGTGATGTCTGTGTTTACATGACCTGTAGCTGTAGTATATGCGAAGTGT 847
Db 804 GTCCGCTGCGCGGGCTGTACACCGCGCTGCAACGCGCTGAGCTGTACCGGAGGCA 863
Qy 848 TGAAGAGAAATGCACTGCAACAGTTGAAGCTTCACTAGCTTCAATGACCAATTT 907
Db 864 CGACACAGGCGGCGCTCGACAAAGCTCGAAGCTTTCGCGCGGCACTT 923
Qy 908 TTATGGCTCTTCAAGAACTCAATCAAGATTAGTGAAGACGCGCATGGAAGTACC 967
Db 924 CTACGGCTTCCCGGACGCGCGGACGAGCGTGCCTGCGCGGACGTGGAGCTGCC 983
Qy 968 CGAATCCTTTTCTATGATCAGAGATATATATCCATGTTTGTGTGTAATGCTCGA 1027
Db 984 GCGGAGATGACCGCGGCGCGCGCGCGGTGTGCTGCGCGCGGCGGAGGATCGG 1043
Qy 1028 CTGGTTGC 1035
Db 1044 CTGGCGGC 1051

RESULT 13

ACA24115

ID ACA24115 standard; DNA; 1092 BP.

XX ACA24115;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #5772.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX Borrelia cepacia.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342922P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JM;

XX Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX P-PSDB; ABU20245.

PT Isolate candidate molecules for rational drug discovery programs.
XX
XX
PS Claim 14; SEQ ID NO 11985; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1092 BP; 175 A; 396 C; 362 G; 159 T; 0 U; 0 Other;
Query Match 20.8%; Score 264.4; DB 8; Length 1092;
Best Local Similarity 56.2%; Pred. No. 6; e-66;
Matches 540; Conservative 0; Mismatches 411; Indels 9; Gaps 2;
Qy 14 GCTTCATATCAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73
Db 63 GCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 122
Qy 74 GGCAGTTGTCTCTCAGAGTGCATCACTTGGAGGCAATAGTCAATGCCAAATTTGAA 133
Db 123 GCGCGTCTGCGGACACCGCGCGCAATGCGCGCGCATGCTCATGCGCAACTGAA 182
Qy 134 GCTTCATATCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 193
Db 183 GCGCGCTGCTCAGACACCGCGCGCAAGGCTATGCGAGGCAATCTTCGCGCGCT 242
Qy 194 ACCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 253
Db 243 GCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
Qy 254 TATGAAATCAAACTAGCAAGAGAGCAGTGTATTTGGGTTGAAGTTGATCCTTGC 313
Db 303 CGAGAAATCG 362
Qy 314 TGTGTCACGACAAATTTCTCAAGTGAAGTACGTGATTTTTCGGGAAGTTTACAGT 373
Db 363 AGGCGCCACGACGAATTCGACATGGGATACCGATC--TCGGAAATGTGCAAGAC 419
Qy 374 TCTCAAGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
Db 420 GCTGAGCGCATGACGAAACCGCGATCGCTGCTGTGACGCGCGAGTGAACGATGC 479
Qy 434 TGAGTTGACATGTTTGTATGAGAAAGATATTCATGAAACGGTTTAAACGCTTGGT 493
Db 480 GTGATGACCTGTTCGACCGCGAAGGTGTTCATGACCGCGGTGATGAGACCGCGCTGCG 539


```
QY 494 GCAGAAATTTCCACAAATTGAGTGTGATGAGACATGTTACCACTTGTGCTGTTAA 553
Db 540 CCGGAGATTTCCGGGAGCCCTGAAAGTGTGTTGAAACATCAACGAGAGACGGGCGGA 599
QY 554 GTTTGT-----TGAATCTTGTGACTGAAAGATTGTTGACAGCACTGTACCCCAACA 607
Db 600 CTACGTCCGCGATGCCGACGGCGGCCCTGCTCGGGCGACGATCAACCGGCACTCA 659
QY 608 TCTGTTTGAACAGCAATTTCTCTTCCAAAGGGGCTTACACCGCATATTAATGCTCT 667
Db 660 CCTCTGTAACAACGCAACGCGCTGTTCTGCGGGGATCGCCCGCATTAATCTCTCT 719
QY 668 TCCAGTCTCTCAAAAGAGATCCACAGGAGGACTTGTGTCAAGCTGTAACAAGTGAAG 727
Db 720 GCGGCTCTGAAGGCGGAGACGCAATCCGCTGCGCTGTGAGAGCCGCGAGCTCGGCA 779
QY 728 TAAAGATTTTCTTGGGACTGATGCTCTCTCATGATGACGAAGAAAGAGTGTTC 787
Db 780 CCGCGCTCTCTCTCCGCGACCGACGCGCGCGCGACGCGCGCAAGGAAACGCG 839
QY 788 TTGTGATGTGTGTGATTTTACAATGCACTGTGACCTTGTGCTGATATATCGAAGTGT 847
Db 840 GTGCGTTTGCGGGGCTGCTACACGCGCTGCAACGCTGATGACCGGAAAGGTT 899
QY 848 TGAAGAGAAATGCACTGACAACTTGAAGCTTCACTAGCTTCAATGACCAATTT 907
Db 900 CGACAGCGCGGGCGGCTGACAAAGCTGGAAGGTTGCGAGCTTCTTGGCGCCGATTT 959
QY 908 TTATGAGCTTCTTGAACAACTCAAAATTAAGTTAAGTGAAGCGCATGGAAGTACC 967
Db 960 CTACGGGCTCGCGCGACGCGCGAGCGGTCAAGCTGCGCGGAGCGCTGGGAACCTGCC 1019
```

RESULT 14

ABA92787_3

Continuation (4 of 7) of ABA92787 from base 300001 (Buchnera sp. genomic DNA SEQ ID NO:1)

WP Sequence Split info 7 fragments LOCUS ABA92787 Accession ABA92787

WP	Fragment Name	Begin	End
WP	ABA92787_0	1	110000
WP	ABA92787_1	100001	210000
WP	ABA92787_2	200001	310000
WP	ABA92787_3	300001	410000
WP	ABA92787_4	400001	510000
WP	ABA92787_5	500001	610000
WP	ABA92787_6	600001	640681

Query Match 20.84; Score 264.2; DB 6; Length 110000;

Best Local Similarity 54.04; Pred. No. 6; e-65;

Matches 564; Conservative 0; Mismatches 478; Indels 3; Gaps 1;

```
QY 20 AATCACAACAACCTGATGATGTCATCTCCGTGATGATGTTCTTAAGCAGT 79
Db 68078 AATCATTAATAACGTATGATGGCATGTTCATTTAAGACAAATGATTTTAATCAAT 68137
QY 80 TGTCTCTCACAGTGCATCTCTTGGAGGGCAATAGTCATGCCAAATTTGAAGCTCC 139
Db 68138 TATTAAGTATACGTGTAATTTTATTAAGAGCTGTATATATGCCAAATCTTAATAGTCC 68197
QY 140 TATCACTACCACTGCTGCTGTGATGATACCGGGAGGGGATATTGAATCTTACTCT 199
Db 68198 GATTAAGATGTTGTTGAAAAAGCATTCGTAATGTAATTTTAAATCAATGCACTT 68257
QY 200 TGATAGTGAATTCACCTCTTATGACATTTATTTGACAGATACCAACGATCTATGGA 259
Db 68258 GAATTAATTAATCAACCGTATATGACTTGTATTTAATCAATTCACAAAGCCCTAAGA 68317
QY 260 AATCAACATGACGAAGAGAGCCAGTCTGATTTGGGGTGAAGTTGTACCTGCTGTGTC 319
Db 68318 ACTAGATTTGTTTCTTAAAAAATATTTGAGAGCTTAATCTTATCCAAATGCTTG 68377
QY 320 CACGACAAATTCACAGATGAGTGAATCTTTTCCGAGAGTGTTCACAGTTCTACA 379
```

```
Db 68378 CACGACAAATTCAAAACTGTATAAAAAAAT---ACTGATATTACTCTGTTTGA 68434
QY 380 AGAATGTTGAGCATATATATGCTCTGCTGTGTTGAGAGAGTTACTATCTGAGGT 439
Db 68435 ATGTATGAAAAAATATGAAATGCAATTAATAATTCATGTGTGAAGATTAACAAAAATAT 68494
QY 440 TGACATGTTGATGAGAAAGATATCTTGAACAGTTCTTAAGACCTTGGTGACAA 499
Db 68495 TGATATCTATGATGAGAAAGCAAAATTTATTAATAAAAAAGTTAATCTTTAAGAAAAA 68554
QY 500 ATTTCCAAATTTGAAGTGTGATGAGCATGTACACCATTTGATGCTTTAAGTTGT 559
Db 68555 ATTTCCAAAGTTAAGATATGATTTGAACATTTACACCAAGAAATCTGTGAATATAT 68614
QY 560 TGAATCTTGCATGAAAGATTGTTGACAACTGTACCCCAACAATCTTGTGTTGA 619
Db 68615 CAAATAATACGATGTTAATTAATCTTATCTGCAACTTTACACCAATCATTTAATCTAA 68674
QY 620 CAGAAATCTCTCTTCCAAAGGGGCTTACACCGCATTAATTAATCTGCTTCCAGTCTCA 679
Db 68675 TCGAATGATATGTTTATGATGAGATTCACCAATATCTTTAATCTGATTTTAA 68734
QY 680 AAGAGATCCAACAGGAGCACTTGTGTCACTGTAAACAAGTGAAGTAAAGATTTT 739
Db 68735 AAAAAACAAACATCGAATGCACTAAGAAAGCCATCTCTAATGAGATTAACATTTTT 68794
QY 740 TCTTGGAGTGAATGCTCTCTCATGATGACGAAGAAAGAGTCTTGTGATGTGC 799
Db 68795 TTTAGGAACGATACAGCTCCACATCTTCAATTAATTAATTAATGCTTGGATGTGC 68854
QY 800 TGTATTTTCAATGACCTGTGACCTGTGATGATATGACGAAGGTGTTGAAGAAAGAAA 859
Db 68855 GGTATATTTATATGCTCATCATCTCTATTAATCTTATGTTAAGTATTTGAGAAATGAG 68914
QY 860 TGCATCTCGAACAAGCTTGAAGCACTCACTCACTCAATGAGACAGATTTTATGAGCTTC 919
Db 68915 ACCGTTAAATATTTTACATCTTTTGTCTGAAATATGCTCTAATTTTATTAACATGCC 68974
QY 920 TGAAGAACAACTCAAGATTAAGTTAGTGAAGCGGCATGGAAGTAACTCCATCTTTC 979
Db 68975 AATTAAACAAGAAACATTAATTAATCAAAACCGTGTAAATTAATTAATAAATAATA 69034
QY 980 TTATCATCAGAGATTAATTAATTTCCCATGTTTGTGCTGTAATAATGCTGATGCTGCG 1039
Db 69035 TGTGTGAAGAAATGATTAATTTCAATTTTATTCAGGTGAATTTTAACTGCTCAATTGA 69094
QY 1040 TCCCTCTGGAATTAATTTGATTT 1064
Db 69095 AAGTATTAATAATCAAGATATTT 69119
```

RESULT 15

ACAL9166

ID ACAL9166 standard; DNA, 1047 BP.

XX ACAL9166;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #823.

XX Antisense; ds; prokaryotic essential gene; cell proliferation;

XX drug design; gene.

OS Escherichia coli.

PN W0200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342933P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR P-PSDB; ABU15296.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14, SEQ ID NO 7036; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1047 BP; 251 A; 292 C; 264 G; 240 T; 0 U; 0 Other;
 Query Match 19.3%; Score 245.6; DB 8; Length 1047;
 Best Local Similarity 54.1%; Pred. No. 1.9e-60;
 Matches 547; Conservative 0; Mismatches 459; Indels 6; Gaps 2;

QY 261 ATCAACTAGCAAGAGAGACGAGTGGTATTTGGGGTGAAGTTGACCTGCTGTC 320
 DB 268 CTGAGCGGGGATTTAACAAGCGGTTCACCGCTGCAAACTTAAACCGGCAACGCA 327
 QY 321 ACACAAATTCACAGATGAGTGAATCTTTTGGGAAGTTTACCAGTTCTACAA 380
 DB 328 ACCACTAATCCAGCCAGCGGTGACGTCAATTTAGCGCAA--TCATGCGGTACTTGG 384
 QY 381 GAAATGTTGACATATATGCTCTGCTGCTTCAATGAGAGTTCTAATCTGAGGT 440
 DB 385 CGCATGGAATAATCGTATGCGCTACTGTGTGATGATGATGACATGCAATATC 444
 QY 441 GACATGTTGATGAGAAAAAGTATTCATTGAAACGGTTCTAAGCCGTGTGCAAAA 500
 DB 445 GACATTTTGTATGTGAAGCGCTTATTAAGAAAGGTATGAACTCTGCGCAGCGC 504
 QY 501 TTTCACAAATTGAAGGTGATGATGAGATGTTAACCACTTATGCTGTTAAATTGTT 560
 DB 505 CTGACTGCGTGAAAGTCGTTTGTGAGCATATCACCACCAAGATGCTGCCGCTATGTC 564
 QY 561 GAATCTTGCACTGAAGATTTGTCAGACAACTGTACCCCAACAACCTTTTGAAC 620
 DB 565 CGTACGGAATGAAAG--GCTGGCTGCACATCATCTCGACACTTGATGTTTAAAC 621
 QY 621 AGGAATTCCTCTTCCAAAGGGGCTTACAAACCGCATTTACTGCTTCCAGTCTTCAA 680
 DB 622 CGAACCATATGCTGTGAGAGCGCTCCGCACTGATTTGTCTAACCATCTCTCAA 681
 QY 681 AGAAGATTCACAGAGGAGGACCTTGTGCTGAGCTGTAAACAAGTGAATTAAGATTTT 740
 DB 682 CGTAATTTTACCAACAGGCGATTCGTAATGCTGTCGCCAGCGGTTTAAATCGATATTC 741
 QY 741 CTGGGACTATATGCTCTCTCATGATAGACGAAGAAAGAGTTCTTGATGATGCT 800
 DB 742 CTGGGTACGATTTCTGGCCACATGACGTCATTCGCAAGAGACAGTGGCGTGC 801
 QY 801 GGATTTTACAAATGACCTGTAGCTTGTACATATATGCAAGAGTGTGAAAAGGAAAT 860
 DB 802 GGCTGCTTCAACGCCCAACCGCGCTGGCATTTACCTACCTTTGAAGAAATGAT 861
 QY 861 GCACTTCACAAAGCTTGAAGATTCATGCTTCAATGAGCAAGATTTTATGCGCTTCT 920
 DB 862 GCTTTGAGACACTTGAAGATTCGTCTGTAAACGGCCGACGTTCTATGCGTTC 921
 QY 921 AGAACAATCTMAAGATTTAAGTTGATGATGAAGCCATGAAAGTACCCGAATCTTTTCT 980
 DB 922 GTCAACGACACATTCATCGAATGCTGACGTGAAGAGCAACAGGTTGCTGAAGACATGCA 981
 QY 981 TATGATCAGAGATATTTATCCATGTTGCTGTGAAATGCTCGACTGCT 1032
 DB 982 CTGACTGATGACACGCTGTGTGCAATTCCTCGCCGGGGAACGTTACGCTGT 1033

Search completed: February 12, 2005, 02:16:30
 Job time : 774 secs

This Page Blank (uspto)

SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 20.8%; Score 264.2; DB 4; Length 640681;
Best Local Similarity 54.0%; Pred. No. 6,6e-71;
Matches 564; Conservative 0; Mismatches 478; Indels 3; Gaps 1;

```
20 AATCAACAACCGATGATGGCATCTTCATCCCGATGAGTGGTCTTAAAGCAGT 79
   |||
Db 368078 AATCATTAACCTGATGATGGCATGTTCAATTAAAGACAATGAGATTTTAAATCAAGT 368137
   |||

80 TGTCTCTACAGTGCATCATCTTTGGAGGCAATGTCATGCCAAATTTGAAGCTCC 139
   |||
Db 368138 TATTAAGTATCTGTAATTTTATTAAGAGCTGTTATATGCCAAATCTTATATAGTCC 368197
   |||

140 TATCACTACCACTGCTGCTGCTGTAGCAATCCGGAGGCGAATTTGAATCTTACCTGT 199
   |||
Db 368198 GATTACGAGTGTGTTGAAAAGCAATGGCTATCGTAATAGAAATTTTAAATCAATGCACCTT 368257
   |||

200 TGATAGATGATTTCAACCTCTTATGACCTTATTTGACAGATACACGACCTGATNGA 259
   |||
Db 368258 GAATTTTAAATTTCAACCGTTAATGACTGTGTTTAACTTAATTCACAGACCTTAAGA 368317
   |||

260 AATCAACTAGCAAGAGAGAGACCGAGTGTATTTGGGGTGAAGTTGATCCCTGCTGTC 319
   |||
Db 368318 ACTAGAAATTTGTTTTCTTAAABAAATTTGTAAGCAAGCTAAATTTCTAACAATGGTTG 368377
   |||

320 CACGACAAATTTCTCAGATGAGTGAAGTACTGATCTTTTGGGAGTGTTTACAGTTCTACA 379
   |||
Db 368378 CACGACAAATTTCAAAACTGGTATATAAAAAAATTT--AGTATATTTCTCTGTTTGA 368434
   |||

380 AGAAATGTTGAGATTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 439
   |||
Db 368435 ATGTATGAAAAAATAGAAATGCCATTAATTAATTAATTAATTAATTAATTAATTAAT 368494
   |||

440 TGACATGTTTGAATGAGAAAAAGATTCATTTGAACGGTTCTAAGACCGTTGTGTGACAA 499
   |||
Db 368495 TGATATCTATGATGAGAGAAATTTATTTGAAAAAAGCTTATGATCTTTTAAAGAAAAA 368554
   |||

500 ATTTCCAAATTTGAGGTCGTGATGAGCATGTTTACCAACATGATGCTGTTAAGTTGT 559
   |||
Db 368555 ATTTCCAAATTTAAGATAGATTAAGAAATATTAACAACCAAGATCTGTAAGATTAAT 368614
   |||

560 TGAATCTTGAAGTGAAGATTTGTTGACGACACTGTCAACCCACAAATCTGTTTGA 619
   |||
Db 368615 CAAAAAATAACGATGTATTAATTAATCTTAATCTGAACATAATTAACACACATCAATTAATGCTAAA 368674
   |||

620 CAGAAATTCCTCTCTTCCAAAGGGGGCTTAACACCGCATTAATTAAGCTCTCAAGTCTCA 679
   |||
Db 368675 TCGAAATGATATGTTTATGTTGGGATTTCAACCATATTTTACTGTTTACGATTTTAA 368734
   |||

680 AAGAGATCCACAGGAGGACCTTGTGTCACTGTAAACAAGTGAAGTAAAGATTTT 739
   |||
Db 368735 AAAAAACAACATGAGATGGAAGTAAAGAAAGCCATCTTAATGAGATTAACATTTT 368794
   |||

740 TCTTGGAGCTATATGCTCTCTCATGATAGACGAAGAAAGAGTCTTGTGATGTGC 799
   |||
Db 368795 TTTTGGAGAGCATATGAGTCTCAACATCTTCAATAAATTAATTAATGCTTGAATGTGC 368854
   |||

800 TGTATTTTACAAATGACCTGTAGCTGTGATGATGATGATGATGATGATGATGATGATGAT 859
   |||
Db 368855 GGGTATATTTATGCTCATCATCTCTATTAATCTTAAAGATTTGAGAAATGAG 368914
   |||

860 TGACATCGACAGCTTGAAGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 919
   |||
Db 368915 AGCGTTAAATATTTAATCAATCTTTTGTCTGAAGAAATGCTCTAAATTTTAAATGATGCC 368974
   |||

920 TAGGAACAATCAAGATTTAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 979
   |||
```

```
Db 368975 AATTAACAAGAAACCATTAACATTAATCAAAAAACCGTGTAAATTTTAAAAATTA 369034
   |||
Qy 980 TTATGATCAGAGATATTAATCCCATGTTTGTGTGTAATGCTGAGCTGTGCGGC 1039
   |||
Db 369035 TGTGGAGAAATGATGATTAATTCATTTATCAGGTGAATTTTAACTGGTCAATTGA 369094
   |||

1040 TCTCTCTGAGATCATTTGTCAAT 1064
   |||
Db 369095 AAGTATTAATTAACAGATATTT 369119
   |||
```

RESULT 4

```
US-09-540-236-410
; Sequence 410, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT.
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 410
; LENGTH: 1068
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-410
```

Query Match 19.1%; Score 243; DB 4; Length 1068;
Best Local Similarity 55.3%; Pred. No. 7e-66;
Matches 552; Conservative 0; Mismatches 420; Indels 27; Gaps 3;

```
Qy 6 AAAATGAGCTCTCAATACACAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 65
   |||
Db 7 ACAAGCATTAATCTCATATTCAGCAGAGAGTGGCATTCATCTGCTGATGATGAA 66
   |||

66 GTTCTTAAGCAGTGTCTCTCAAGTGCATCACTTTGGAGGCAATAGTATGCCA 125
   |||
Db 67 GCTCTGCAACACCGTTCGGATGCGGCAACAGTTTAAACGTGTGATTTGATGCTT 126
   |||

126 AATTTGAAGCTCTCATACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185
   |||
Db 127 AATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 186
   |||

186 AAATCTTAC-----CTGTGATGATGATTTCAACCTCTTATG 224
   |||
Db 187 AAGCATTAAGAAGACGATTTTACAGCCATCGCAAGATAGTTTGCATCCAGCATG 246
   |||

225 ACACCTTTATTTGACAGATACACCAAGCTCTATGAAATCAAACTAGAGAGAGGCGAG 284
   |||
Db 247 GTGCTTTATTTGACCGATCAAAACCGCAAGATATGATGATGATGATGATGATGATGATGATGAT 306
   |||

285 GTGCTTTATTTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 344
   |||
Db 307 ATTTATCGGAGATTAAGCTGTATCTGCTGAGCAACCAATTTAGCTGATGATGATGATGAT 366
   |||

345 ACTGATCTTTTTCGGGAAGTGTTCACAGTTCTCAAGAAATGTTGAGCATATATGCTT 404
   |||
Db 367 ACTGATTTTTCGGACCAAGCTT---GGTATTTGAGCCATGCAAAAGCATGCGATCCG 423
   |||

405 CTGCTGCTTATGAGAGGTTTACATATCTGAGGTTTACATGATGATGATGATGATGATGATGATGATGAT 464
   |||
Db 424 CTACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483
   |||

465 TTCAATGAACGCTTTCAAGACCGTTGGTGGCAAAATTTCCACATTTGAAGTGTGTGATG 524
   |||
Db 484 TTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
   |||

525 GAGCATGTTACCAACATTTGATGCTGTTAAGTTTGTGATTTTGCATGAGAGATTTGT 584
   |||
Db 544 GAGCATATCACTACCGCGAGTGCAGCGGATTTTGTCTGCT---CAAGGCAATCAATC 600
   |||
```

QY 585 GCAGCACTGTCAACCCCAACAACATCTTGTGTTTGAACGAATTCTCTTCCAAAGGGGC 644
 | | | | |
 Db 601 GCACGCGCAATTACCCCGAGCAATCTTTATTATATGCATATCTGCTGGTGGCGGT 660
 | | | | |
 QY 645 TTACAAACCGATATATTCTGCGCTTCCAGTCTCTCAAAAGAGATGCCAGAGGAGCACTT 704
 | | | | |
 Db 661 ATTAAGCGGCACTATTATTGTTGGCCATTTTAAAGGTCCTGCACACCAAAAGCGTTTG 720
 | | | | |
 QY 705 GTGTCAAGCTTAACAAGTGGAAGTAAAGAATTTTTCTTGAGCTAGTAGTGTCTTCAT 764
 | | | | |
 Db 721 CTTTGAAGTGGCAACCGTGGCAATCCTAAGTTTTCTTAGGAGCAAGTTTCGGCCCTCAAT 780
 | | | | |
 QY 765 GATAGACGAAGAAAAGATGTTCTTGTGATGTGCTGTAATTACATGACCTGTAGGCC 824
 | | | | |
 Db 781 GCCACACACACGAAAGAAATCATGCTGTGCTGTGCAGTGTGTTAATGTCGCTTGACGCT 840
 | | | | |
 QY 825 TTGTCAGTATATGCGAAGGTGTTTGAAGAAAGAAATGCCTGACGACAGCTTGAAGATTC 884
 | | | | |
 Db 841 CTACACCTCTACGCCATGGCAATTTGAAATGTCCAAAGCATCTGATTAACCTTAAAAACCTT 900
 | | | | |
 QY 885 ACTACCTTCATATGACACGATTTTATATGAGCTTCTAGAGACAATCTCAAAAGTTAAGTTG 944
 | | | | |
 Db 901 GCACCCGATATGAGGACAAATTTATGCTTAGCAATCAACACACTCAAAATCACCTTA 960
 | | | | |
 QY 945 AGTAAGACGCATATGAGAGTATCCCGAATCCTTTTCTTAT 983
 | | | | |
 Db 961 ATTAAGCAACCAAGACCAATACCTGAGAGTTTCAATAT 999

```

RESULT 5
US-09-252-991A-920
; Sequence 920, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.116
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 920
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-920

```

Query Match	19.1%;	Score 242.6;	DB 4;	Length 1197;
Best Local Similarity	53.8%;	Pred. No. 1e-65;		
Matches 547; Conservative	0;	Mismatches 464;	Indels 6.	Gap 2

QY 15 CTCTCAATCAACAACCTGATGATTTGGCATCTTCATCTCCGTGATGTGATGTTCTTAAAG 74
 Db 163 CTCACCCCTCCTGGCCCCCGAGACTGGCAATTCACCTGGTGACGGTCCCGCGCTGGCC 222
 QY 75 GCATTGTCTCTCAGAGTGCATCACTTTGGAGGGCAATAGTATGCGCAAAATTGGAAG 134
 Db 223 AATACCGTGGCGGACGGCGCCCGACCTTGGCGCGGCATGTATCGCCAACTGGTGT 282
 QY 135 CCTCTATCACTACCACTGCTGTGCTGTAAGCATACCGGAGAGCAATTTGAATCTTTTA 194
 Db 283 CCGCGGGTGGCAACGCCGCCGAGCCGCTTACCGGCAACGGCATCTCGCGCGCCGC 342
 QY 195 CCTGTTATATGTAATTTCAACCTCTTATAGACATTTATTTGACAGATACCAACCAAGTCT 254
 Db 343 CCGCGCGCCGCGCTTTCGAGCGCTGATGATGTGTACCTCACCGAACCGGACCAAGGCC 402
 QY 255 ATGGAATCAAACTACGACGAGAGGCCAGTGTGATTTGGGGTGAAAGTTGATCCCTGCT 314

Db	403	GAGAAATCCGATCGGCCAAGGCCAGCGGCTTCGTCATGCGCCAAAGCTCATCCGGCA	462
QY	315	GGTGCCAGACAATTTCTCAGATGAGTGAATCTTTTCGGGAAGTGTTCACAGTT	374
Db	463	GGGGCACCAACCACTCGGATTTCTGGCGTAC--CCGATCGACAATATCTTCGAACG	519
QY	375	CTACAGAAATGGTTGAGCATATATGCTCTCTGCTGTTCAATGAGAGTTAACTAATCT	434
Db	520	CTCGAGCGCATGGCCGAGGTGCGCATGCGCTGCTGGTGTGATGGGAGGTGACCGCGCC	579
QY	435	GAGGTTGACATGTTGATAGAAAAGGTAATTCATTGAAACGGTCTTAAGACCGTGGTG	494
Db	580	GAGGTGACGATTCGATCTGAGAGACGATTCATCAGACGACCTGCGCCGGGTGTC	639
QY	495	CAGAAATTTCCACAATGAAGGTGATGAGACATGTTACACCATGATGCTGTTAAG	554
Db	640	GAGCGCTTCCGACCCCTGAAAGTGTCTTCAGACACATCAACCGCGACGCGCCGACG	699
QY	555	TTTGTTGAATCTTGACACTGAAAGATTTGTTGACAGCACTGTGCACCCCAACAATCTTGT	614
Db	700	TTCTTCGCGGAAGCCCC--GGCCAAGTGGGGCGGACCATTAACGCCCATCACTGCTG	756
QY	615	TTGAACAGGAATTCCTCTTCCAAAGGGGCTTACAAACCGATATTAATGCTCTTCAGTC	674
Db	757	TACAAACGCAACACATGCTGTGTGGGGATCCGTCCGACCTTATATCTGCGCATC	816
QY	675	CTCAAAAGAGATCCACAGGGAGGCACTGTGTCAAGCTGTAAACAAGTGAATTAAGA	734
Db	817	CTCAAGGCAACACCAACGAAAGCCGTGTGAAGCCGCGGTGACGGCAATCCGAAA	876
QY	735	TTTTTTCTGGGACTGATAGTGTCTCTCATGATAGCGAAGAAAGATGTTCTTGTGA	794
Db	877	TTCTTCCTCGGACCACTCGGGCGGACAGCCCGCAAGCCAAAGAACCCCTTCGGGT	936
QY	795	TGTCGTGATTTTACAATGCACTGTAGCTTGTCAGTATATGCGAAGGTGTTGAAGA	854
Db	937	TGGCGGGGTGTACAGCCCTACGCCGCGCATCGAGCTATAGCCGAGGCTTCGAACG	996
QY	855	GAATAATGCATCGACAGCTTGAAGATTCATTAATCTTCATGACACAGATTTTATGGG	914
Db	997	CGCAATGGCTGACAGCTGAAAGGCTTCGCAACTGCAAGGCGCCGACCTTCAATGAC	1056
QY	915	CTTCTGAGAACAACTCAAGATTAAGTTAGTAAAGACGCCATGGAAGTATCCGATCC	974
Db	1057	CTGCGCGCAACACGACCGGATCACTGTGTCGCGAGGAATGACAGCGCCCGCCAGC	1116
QY	975	TTTTCTTAATGCACTCAGAGATTAATTCACAGTTTGTGTGAATGCTGACTGG	1031
Db	1117	CTCCCTTCGGGAATTCGACTGTGTGCGCTGCGCGCGCGGACGACCTTCGCGTG	1173

```

RESULT 6
US-09-252-991A-1041/c
; Sequence 1041, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1041
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-1041

```


QY	645	TTACACCGCATATATTACTGCTTCAGTCCCTCAAAAGAGATCCACAGGGAGGCATT	704
Db	72000	ATTAAAGCGGCACATATTATTGTTCGCCATTTTAAAGGTGCTCACAACCAAAAGCGTTGG	71941
QY	705	GTGTGAGTGTAAACAAGTGGAAAGTAAAGATTTTTCCTTGGGACTGATAGTGCCTCAT	764
Db	71940	CTTGAAATGGGCACACAGTGGCGCAATCTTAAGTTTTTCTTAAAGGACAGATTCCGCCCTCAT	71881
QY	765	GATAGACGAGAAAAAGAGTGTCTTGGATGTGCGGTATTTACAACTGACCTTAGCC	824
Db	71880	GCCACACACAGAAAGATCATCGTGTGCTGTGCGAGTTGTTATAGTCCCTTGACACCT	71821
QY	825	TTGTCACTAATATGGAAGTGTGTTGAAAAAGAAATGCATCGACAAAGCTTGAACATTCC	884
Db	71820	CTACCACTCTAAGCGCATGTGCATTTGAAAGTGCATAGCATTTGATTAATTGAAACTTT	71761
QY	885	ACTGACTTCATGAGACCCAGATTTTATATGGGCTTCTTAGAACAACCTCAAAGATTAAAGTTG	944
Db	71760	GCCAGCGGATATGTGTGACATTTTATATGGCTTAGCAATCAACACCATCAATCACCCTA	71701
QY	945	AGTAAAGCGGCATGAAGGTACCCGATCCTTTTCTTAT	983
Db	71700	ATTAACAACACACAGACCAATACTAGAGTTTCAATTAT	71662

RESULT 8
 US-09-489-039A-418
 ; Sequence 418, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 418
 ; LENGTH: 1059
 ; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae
 ; US-09-489-039A-418

Query Match	17.4%	Score 221.6	DB 4	Length 1059
Best Local Similarity	54.4%	Pred. No. 3.9e-59		
Matches 490	Conservative 0	Mismatches 404	Indels 6	Gaps 2
Qy	20	AATCACACAACTGATGATGGCATCTTCATCTCCGTATGTGATGTTCTTAAAGCAGT	79	
Db	39	AATCGCGCGCCAGACGACTGCGCATATCTGCGCATACGATATAGCTGAAAAACCGT	98	
Qy	80	TGTCCTCAACGTGACATCACTTTGGAGGGCAATAGTCATGCGCAATTTGAAAGCCGCC	139	
Db	99	CGTGGCTTATACAGTGAATTTTATGGCGGGCGATGTGATGCCAATCTGTGTGCGGCC	158	
Qy	140	TATCACTACCACTGTGCTGCTAGATACATACCGGAGGCGATTTGAAATCTTTACCTGT	199	
Db	159	TGTACACACCGTCGCGGCGCTATCGCTTACCGGACAGCATTTATGACCCCGGCGCGC	218	
Qy	200	TGATATGATTTCAACCCCTTTATGACACTTATTTGACAGATACCAACAGTCCCTATGGA	259	
Db	219	CGGGCAGATTTTACCCCGCTGATGACCTGTGATCTACAGACTCTTTCGACCCAGCGGA	278	
Qy	260	AATCAAACTGACAAAGAGAGCCAGGTCTGATTTGGGGTGAAGTTGTACCTCGTGGTGC	319	
Db	279	GTGGAACCGGGCTTTTAACAGAGGGGTTCACCGCGCAAGCTCTACCCGGCCAAATG	338	
Qy	320	CACGACAAATTTCAAGATGAGTGACTGATCTTTTTCGGGAAAGTTTACCACTTTCA	379	
Db	339	CACCTACTCACTCAGGACGAGCGGTCAACAGACACCGACGCA--TTATGCAAGTGTGGA	395	

QY 380 AAAAAATGGTTAGACAATAATATGCTCTGCTGGTTGATGAGAGGTTACTTAATCTGAGGT 43.9
 Db 396 GCGGATGGAAAAAATCGGGGATGCGGTTGCTCTGTTACACGGCCAAAGTCACTACGCCGAGAT 45.5
 QY 440 TGACATGTTTGTATGACAAAAAGGATTTCAATGAAACGGTTCTMAAGACCGTTGGTGCAGAA 49.9
 Db 456 TGAATATTTTGCATGCTGTAAGACCGCGCTTTATTBAGACGGTATGAGACCCCTGCGCCAGCG 51.5
 QY 500 ATTTCACAATATGAAGGTGCTGATGAGAGATTTACCAACATGATGCTGTAAATTTGT 55.9
 Db 516 CCTGCGCGGCTGAAGATGGTGTATTTGAAATATATACCAACCAAAAGCCGCCCAAGTACG 57.5
 QY 560 TGAATCTTGCATGAAAGATTGTTGTCAGCAACTGTCAACCCCAACAATGTTGTTTGA 61.9
 Db 576 ---GCGTGAAGGTAATGAACCTGCTGGTGTACCATCAACCCCGAGCATGTATGTTCA 63.2
 QY 620 CAGGATTTCTCTTCCCAAGGGGGCTTACAAACCGATATTAATCTGCTTCCAGTCTCAA 67.9
 Db 633 CCGGAACCAATCTCTGTGGGCGGATATGCGCTTACCTTACTGTCTGCGGTACTTAA 69.2
 QY 680 AAGAGAGATCCACAGGGAGGACCTTGTGTCACTGTATCAAGATGGAAGTAAAGATTTT 73.9
 Db 693 ACGCATATATTCATCAGACAGGCTCTGCGGGAGCTGTGCGCAGGCGCTTACGCCGCGCTT 75.2
 QY 740 TCTTGGGACGTATAGTGTCTCCATGANTATGACGAAAGAAAGAGTGTCTTGTGATGTGC 79.9
 Db 753 CTTGCGCACCGACTCCGACCCGACGCCGCAACCGTAAGSAAAGCCAGCTGTGGCTGGCG 81.2
 QY 800 TGTATATTTACAAATGCAACTGTAGCTTTGTGATATATGGAAGAGTGTGTAAGAAAGAA 85.9
 Db 813 CGGCTGCTTTAAGCGCCCGACCGGCTCTGGCAGCTACGCCACGGTGTGAAAGAAATBA 87.2
 QY 860 TGCATTCGACAAGCTTGAAGCAATTCATCACTCACTCAATGACCAAGATTTTATGCGCTTC 91.9
 Db 873 CGCTCTGACAGACTTTGAGGCTTTTCTGCTGCTTAAACGCGCCGCTTCTACGGTCTGCGC 93.2

RESULT 9
 US-09-252-991A-1008
 ; Sequence 1008, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 1008
 ; LENGTH: 1266
 ;
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-1008

	Query Match	10.2%	Score 129.8	DB 4	Length 1266
	Best Local Similarity	52.9%	Pred. No. 4e-30		
	Matches 303	Conservative 0	Mismatches 267	Indels 3	Gaps 1
QY	459 AAGGATTCATGAAACGGTCTTAAGACGGTGGTCAGAAATTTCACAAATTGAAGTC	518			
Db	2 AAGCAGTTCATGACAGACGACCTGCGCGGAGTGTGACAGCGCTCCGACCCCTGAAGGTG	61			
QY	519 GTGATGAGACATGTTACCACTTATGCTGTTAATTGTTGAATCTTGCACTGAAGGA	578			
Db	62 GCTTCGAGCACAATCACACCGCGAGCGCGCCAGTTGTGCCGGAAGCCCC--GCC	118			
QY	579 TTGTTGACAGACATGTCACCCCAACAACATCTTTGTTTGAACAGGAATTCCTCTTCCAA	638			

Db 119 AACGTGGCGGACCATTAACGCCCATCTGCTGTACCAACCGCAACATGCTGTC 178
Qy 639 GGGGGCTTACAACCGCATTAATTACTGCTTCAGCTCTCAAAAAGAGATCCACAGGAG 698
Db 179 GGGGATATCCGTCGCGCACTTCTATTGCTCCGATCTCAAGCGCAACCCACAGGAA 238
Qy 699 GCACTTGTGTACGTGTAAACAGTGAAGTAAAGATTTTCTTGGGATGTATAGTCT 758
Db 239 GCCCTGTGTAGCCCGCGGTGAGGGGCAATCCGAAATCTCTCCGCGACGCTCGCG 298
Qy 759 CCTCATGATAGAGAAAGAAAGTGTCTTGTGATGTGCTGTATTAATGACACT 818
Db 299 CCGGAGCGCCGCGCAAGGAAAGCCGCTGCGTTCCGCGGCTCAAGGCTTAC 358
Qy 819 GTAGCTTGTAGTATATGCGAAGGTGTGAAAGAAAGAAATGCACTGCAAGCTTGA 878
Db 359 GCGGCAATGAGCTGTATGCGAGGCTTGGAAACAGGCAATGGCGTGGCAAGCTGAA 418
Qy 879 GCATTCACTAGCTTCAATGAGCAAGATTTTATGGGCTTCTTGAACAACTCAAGAT 938
Db 419 GGGTCCGCAAGCTGACGCGCCGACTTCTATGCTGCGCGCAACACCGAGATC 478
Qy 939 AAGTTGATAGAGCCGATGAGATCCGAAATCCTTTCTTATGATCAGAGATATT 998
Db 479 ACCCTGTGCGCGAAGATGCGAGCGCCGCGCACTCCCTTCGCGCACTTGACATG 538
Qy 999 ATTCCAGTTTGTCTGGAATGCTGCACTGG 1031
Db 539 GTGCGCTGCGCGCGCGAGACACTCCGCTGG 571

RESULT 10

US-09-252-991A-1117/c
; Sequence 1117, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1117
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1117

Query Match 6.3%; Score 79.6; DB 4; Length 261;

Best Local Similarity 57.1%; Pred. No. 1.1e-14;
Matches 145; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 676 TCAAAAGAGATCCACAGGAGGACTTGTCTACGCTGTAAACAGTGAAGTAAAGAT 735
Db 261 TCAAGCGCAACCCACCAAGAGAGCCGTGTGAGCGCGGTGCGCAATCCGAAAT 202
Qy 736 TTTTCTGGGACTGATAGTGTCTCTCATGATAGACGAAGAAAGTGTCTTGTGAT 795
Db 201 TCTTCTGGGCACTGCTGCGCGCGCGAGCGCCGCGCAAGCAAGCCGCTGCGGTT 142
Qy 796 GTGCTGTATTATTAATGACACCTGTAGCCCTTGTGATATGCGAAGGTGTTGAAAG 855
Db 141 GCGCGGGGTGCTACAGCGGCTACGCGCATGAGCTGATGCTCAGAGGCTTGAACAGC 82
Qy 856 AAAATGCACTGACAAAGCTTGAAGATTCATGCTTCAATGAGCAACATTTTGTGGC 915
Db 81 GCAATGCGCTGAGCAAGCTGGAAGGCTTGCAGGCTGACGCGCGGACTTCTATGCGC 22

Qy 916 TTCTTGAACAAC 929
Db 21 TGCGCGCAACACC 8

RESULT 11

US-09-252-991A-963
; Sequence 963, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 963
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-963

Query Match 6.2%; Score 79; DB 4; Length 615;
Best Local Similarity 53.4%; Pred. No. 2.7e-14;
Matches 166; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 721 GTGAGTAAAGATTTTCTTGGGACTGATAGTGTCTCTCATGATAGCAAGAAAG 780
Db 1 GCGGCAATCCGAAATCTCTCTCGGCACTGCGGCGGCGCAAGCCGCGCAAGC 60
Qy 781 AGTGTCTTGTGATGCTGTATTAATGACACTGTAGCTTGTCAATATGCGA 840
Db 61 AAGCGCTCGGTGGCGGGGTGCTACAGCCCTTACGCCCATGAGTGTATGCCG 120
Qy 841 AGGTGTTGAAAGAAAGTCACTCGCAAGCTTGAAGATTCATGACTTCAATGAGC 900
Db 121 AGGCTTCAAGACCGCAATGCGTGAACAAGCTTGAAGAGCTTGCACGCTGACGGC 180
Qy 901 CAGATTTTATGCTTCTTGAACAATCTCAAGATTAAGTATGATGAAGCCGATG 960
Db 181 CGACTTCTATGCTGCTCGCGCAACCGACCGATCACCTGTCCGCGAGGAATGGC 240
Qy 961 AGTACCGCAATCTTTCTTATGATCAGGAGATATATCCATGTTTCTGTGAAA 1020
Db 241 AGGCGCGCGCGAGCTCCCTTCGCGCATTTGAGCTGTGCTGCGCGCGCGAG 300
Qy 1021 TGCTGACTGG 1031
Db 301 CACTCCGCTGG 311

RESULT 12

US-09-252-991A-1084/c
; Sequence 1084, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 1084
LENGTH: 294
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1084

Query Match 6.1%; Score 78; DB 4; Length 294;
Best Local Similarity 59.4%; Pred. No. 3.6e-14;
Matches 151; Conservative 0; Mismatches 100; Indels 3; Gaps 1;

QY 306 TACCTGCTGGTGCACGACAAATTTCTCAAGATGAGTGAATCTTTGGGAAAGTGT 365
DB 293 TATCCGCGAGGCCCAACCACTCGATTCTGGCGTGAC---CCGATCGACCAATATC 237
QY 366 TTACCAAGTCTTACAAAGAAATGTTGAGCAATATATGCTTGTGTTGATGAGAGGTT 425
DB 236 TTGGAAGCGCTCGAGCGATGAGCGGAGTGGCATGCGCTGCTGGTGCATGCGAGGTTG 177
QY 426 ACTAATCTGAGGTTGACATGTTGATAGAGAAAGATTTCAATTGAAACGGTTCTAAGA 485
DB 176 ACCCGCGCGAGGTGACGATTCGATCGTAGAGAGAGATTGATGACGACGACCTGGCG 117
QY 486 CCGTGGTGCAGAAATTTCCACAAATGAGGTGCTGATGAGACATGTTACCACTATGAT 545
DB 116 CCGGTGATCGAGGCTTCCGACCTGGAAGTGTCTTTCAGACATCACCGGCGAC 57
QY 546 GCTGTTAAGTTTGT 559
DB 56 GCCGCCCAAGTTCT 43

RESULT 13
US-09-248-796A-4003
Sequence 4003; Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Kelch weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 4003
LENGTH: 1098
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-4003

Query Match 6.1%; Score 77.6; DB 4; Length 1098;
Best Local Similarity 53.2%; Pred. No. 1.1e-13;
Matches 217; Conservative 0; Mismatches 179; Indels 12; Gaps 2;

QY 560 TGAATCTTGACAGAGATTTGTTGAGCACTGTACCCACAAATCTGTTTGA 619
DB 597 TAAACAGAGATTAATTTTGTGTCTCAATCACTGCTCATATTAATTTGAT 656
QY 620 CAGGAATTTCTCTTCCAAAGGGGCTTACACCGCATTAATTAATTCGCTTCACTCTCA 679
DB 657 CATTGAT-----AATGGGCGGAAATCAATTAATTTTGTAAACCAATTCGCA 707
QY 680 AAGAGATTCACAGGAGGCACTTGTGACGTGTACAGTGAAGTGAAGTAAATTTT 739
DB 708 GTTTCAAAAGACAGAAAGCACTGTGTGATCTGCAACTAGTGGAGAACCAATGTTTT 767
QY 740 TCTTGAGACTGATGCTCTCATGATGACGAGAAAGATGT---TCTGTGATG 796
DB 768 CTTTGTCTGATGATGCGCCATCGATTTCATGCAAGCAACTCAGCTGTGTGTG 827

QY 797 TCTGTATTTCAAATGACCTGTAGCCTTGTCAATATGCGAAGTGTGAAAGGA 856
DB 828 TGTGTGTGTTTACCTCAAGACCATGCTATGCTTATGTTGCGAGATTTTGATTAAC 887
QY 857 AATGCACTGACAAAGCTTGAAGCATTCATGCTTCAATGACCAAGATTTTATGGCT 916
DB 888 CACAAATTTGACAAATTTGAGAAATTTGTCAATGATGATGATTTATGATGATT 947
QY 917 TCTTGAACAACCTCAAGATTAAGTTAGTGAAGCGCATGGAAGGT 964
DB 948 ATCAGATGATTTTAAAGAAACATAGATTAATCAGTTGTGTGT 995

RESULT 14
US-09-806-708B-22/c
Sequence 22; Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(1141)
OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAB1 promoters
US-09-806-708B-22

Query Match 4.4%; Score 55.8; DB 4; Length 1141;
Best Local Similarity 12.8%; Pred. No. 8.3e-07;
Matches 120; Conservative 327; Mismatches 483; Indels 6; Gaps 3;

QY 259 AATCAACTAGCAAGAGAGACCGAGCTGTATTTGGGGTGAAGTTTACCTGCTGTG 318
DB 992 RMAAYAAAKAMAGANNMEMYGAAGKWCMAAMATMBWADTAGKWCNNNNNNMTTD 933
QY 319 CCACGACAAATCTCAAGATGAGTACTGATCTTTGCGAAGTGTACCAAGTTCTAC 378
DB 932 VERMAKAKANNNNNNNNYWTACYNRAATNNKATTHMKTTGHAHSKRRTTHRTCRRTX 873
QY 379 AAGAAATGTTGACATATATATGCTTGTGCTGCTGCTGATGAGAGAGTTACTTAATCTGAG 438
DB 872 NNNNNNNARLVYWHHAARRMMAMMTTRNNNNNNNNNNNACRYTRTMABWMSWNNNNN 813
QY 439 TTGACATGTTGATAGAGAAAGTATTTGTTGAAACGGTTTAAAGCCGTTGGTGCAGA 498
DB 812 NNNNNNNNNNTWCHYTTANABBCYRANNNNAAARMATCNYHAAVTTTHTDWCYXTWNN 753
QY 499 AATTTCCAAATGGAAGGCGTATGAGCATGTTACCAACATATGATGCTGTTAAGTTG 558
DB 752 TWYMDMTTMBTTTTRNNMTTSTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 693
QY 559 TTGAATCTTCACTGAAGATTTGTTGACAGACTG--TCACCCCAACAACATCTGTGTTT 616
DB 692 NTDARTNNNTVMBRRMMNTKTRMYSTBRHHYGTATNNNNNNNNNNNNNNNNNNNNNN 633
QY 617 GAACAGAAATTTCTCTTCCAAAGGGGCTTACACCGCATTAATTAATTCGCTTCACTCT 676
DB 632 RMTTKRWTKGGMIVRKVKWRDTCYVDWADSWVYANNRCRDVITYTRNNTYCK 573
QY 677 CAAAAGAGATCCACAGGAGGCACTTGTCAGCTGTAAACAGTGAAGTAAAGATT 736
DB 572 STAHSTYWSNNAMRYRSABNMSSMAATTRNNNNNSGVBVRRAGTMMRRHNNNT 513

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 12, 2005, 05:10:23 ; Search time 811 Seconds
(Without alignments)
9247.207 Million cell updates/sec

Title: US-10-070-277-1

Sequence: 1 ttgcacaaatgagctctca.....tgaagattcactgataaaaa 1271

Scoring table: IDENTITY NUC
Gapop 10_0, Gapext 1.0

Searched: 5378673 seqs, 2950229984 residues

Total number of hits satisfying chosen parameters: 10757346

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications NA:*

- 1: /cgn2_6/ptoddata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptoddata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptoddata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptoddata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptoddata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptoddata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptoddata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptoddata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptoddata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptoddata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptoddata/1/pubpna/US09C_NEW_PUB.seq:*
- 12: /cgn2_6/ptoddata/1/pubpna/US10A_PUBCOMB.seq:*
- 13: /cgn2_6/ptoddata/1/pubpna/US10C_PUBCOMB.seq:*
- 14: /cgn2_6/ptoddata/1/pubpna/US10C_PUBCOMB.seq:*
- 15: /cgn2_6/ptoddata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptoddata/1/pubpna/US10F_PUBCOMB.seq:*
- 17: /cgn2_6/ptoddata/1/pubpna/US10F_PUBCOMB.seq:*
- 18: /cgn2_6/ptoddata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptoddata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptoddata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptoddata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptoddata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	646	50.8	1743	18	US-10-451-554-2
2	607.4	47.8	1826	17	US-10-424-599-6649
3	549.2	43.2	1545	18	US-10-437-963-86830
4	531.6	41.8	1893	18	US-10-435-115-86203
5	345	27.1	1258	18	US-10-425-115-86207
6	311.8	24.5	931	18	US-10-767-701-2040
7	281	22.1	1053	17	US-10-282-122A-41490
8	277.2	21.8	1047	17	US-10-282-122A-42220
9	267.4	21.0	1032	17	US-10-282-122A-8516
10	265.2	20.9	1056	17	US-10-282-122A-14242
11	264.4	20.8	1092	17	US-10-282-122A-11985

12	264.2	20.8	640681	9	US-09-790-988-1	Sequence 1, Appl1
13	245.6	19.3	1047	17	US-10-282-122A-7036	Sequence 7036, Ap
14	241	19.0	1047	17	US-10-282-122A-30305	Sequence 30305, A
15	238.2	18.7	1062	17	US-10-282-122A-15202	Sequence 15202, A
16	238.2	18.7	1062	17	US-10-282-122A-26930	Sequence 26930, A
17	238.2	18.7	92407	17	US-10-672-787-36	Sequence 36, Appl
18	236.2	18.6	1044	17	US-10-282-122A-31411	Sequence 31411, A
19	236	18.6	1032	17	US-10-282-122A-29654	Sequence 29654, A
20	231.2	18.2	1035	17	US-10-282-122A-33544	Sequence 33544, A
21	221.6	17.4	1044	17	US-10-282-122A-33743	Sequence 33743, A
22	219.2	17.2	1047	17	US-10-282-122A-39694	Sequence 39694, A
23	218.4	17.2	1052	17	US-10-282-122A-36568	Sequence 36568, A
24	186.4	14.7	642	18	US-10-021-323-1199	Sequence 1199, Ap
25	183.8	14.5	642	18	US-10-425-115-86208	Sequence 86208, A
26	178.4	14.0	828	18	US-09-732-627A-1146	Sequence 4146, Ap
27	177	13.9	381	11	US-10-424-599-6648	Sequence 6648, Ap
28	166.6	13.1	434	17	US-10-424-599-6648	Sequence 19366, Ap
29	149	11.7	570	17	US-10-021-323-2921	Sequence 2921, Ap
30	138.8	10.9	574	18	US-10-021-323-2921	Sequence 107341, A
31	130.8	10.3	411	17	US-10-424-599-107341	Sequence 6104, Ap
32	129.6	10.2	688	18	US-10-282-122A-18083	Sequence 18083, A
33	114.4	9.0	1008	17	US-10-282-122A-32732	Sequence 32732, A
34	105	8.3	407	17	US-10-282-122A-32560	Sequence 22560, A
35	77.6	6.1	1020	17	US-10-282-122A-32560	Sequence 22560, A
36	67.4	5.3	348	17	US-10-335-977-2162	Sequence 2162, Ap
37	65.2	5.1	1020	17	US-10-335-977-2162	Sequence 2162, Ap
38	65.2	5.1	1113	17	US-10-425-115-89	Sequence 89, Appl
39	58.4	4.6	2580	18	US-10-425-115-89	Sequence 2160, Ap
40	50.4	4.0	438	17	US-10-425-115-164730	Sequence 164730, A
41	48.2	3.8	1992	18	US-10-027-632-250147	Sequence 250147, A
42	42	3.3	652	13	US-10-027-632-250147	Sequence 250147, A
43	42	3.3	652	13	US-10-027-632-250147	Sequence 250147, A
44	42	3.3	652	17	US-10-027-632-250147	Sequence 250147, A
45	42	3.3	652	17	US-10-027-632-250147	Sequence 250147, A

ALIGNMENTS

RESULT 1
US-10-451-554-2
Sequence 2, Application US/10451554
Publication No. US20040111769A1
GENERAL INFORMATION: Rodolphe Arthur Kanhonou
APPLICANT: Ramon Serrano Salom
TITLE OF INVENTION: Sugar beet genes involved in stress tolerance
FILE REFERENCE: 1187-24
CURRENT APPLICATION NUMBER: US/10/451.554
CURRENT FILING DATE: 2003-06-20
PRIOR APPLICATION NUMBER: PCT/EP01/15093
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: EP 00870319.1
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 60/271.656
PRIOR FILING DATE: 2001-02-26
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 1743
TYPE: DNA
ORGANISM: Beta vulgaris
US-10-451-554-2
Query Match 50.8%; Score 646; DB 18; Length 1743;
Best Local Similarity 76.4%; Pred. No. 5.2e-175;
Matches 793; Conservative 0; Mismatches 245; Indels 0; Gaps 0;
Oy 3 GCAAAATGAGCTCTCAATACACAACTGATGATTCGATCTTCATCTCCGATGATGT 62
DB 193 GTAAGATGAGCACTGATCTTACACGCGCTGATGATGCGATCAATCTCCGATGGA 252

RESULT 2
US-10-424-599-6649
; Sequence 6649, Application US/10424599
; Publication No. US20040031072A1

OTHER INFORMATION: Clone ID: PAT_MRT3847_106012C.1
US-10-424-599-6649

Query Match	47.8%;	Score 607.4;	DB 17;	Length 1826
Best Local Similarity	74.3%;	Prod No 730-104		

Matches 778; Conservative 0; Mismatches 268; Indels 1; Gaps 1

3 GCAAAATGAGCTCTCAATCACACAACCTGATGATGGCATCTTCATCTCCGTGATGGT 62

D**b** 360 GTAAAGATGAGCTCACTATTACACAGCTGATGATTGGCATCTTCACCTCCGATGGT 411

63 GATGTTCTTAGGCAGTTGTCTCTACACAGTCACATCACTTTGGAGGGCAATGTCATG 12

Db
420 TCCCTTCTGAGCTGTCTCCCTCACAGTCCGACCATTTGGAGGGCCAGTATG 47

123 CCAATTGAGCCTCTATCATCAACCACTGCTGCTGTAGCATACCGGAGCGGATA 18

Db

480 CCAATTGAAACCACCCATCATCTACCAAGTCGTGCTGCTGCTTATCGAAGCTCCGTTT

Oy 183 TTGAAATCTTACCGTGATGATGCAATTTTCACCCCTCTTAACAAGCCTTTAATTTTGGTCTT

[illegible][illegible]

500 GATGACCTTACCCGAGGAGGAGCCAGGTCGATTTCGGGGTGAG 302

303 TTTTCTTCCCTTGGACCTGTTTATGGTGTGAAC 659

303 1161ACCCCTGCTGTCGACGACAAATTCACAGATGGAGTGACTGATCTTTTCGGGAAG 362

660 TGTATCTGCTGGTGTCTACAACAACCTCCCAANATGTTGTACAGATCTTTTGGAAT 719

363 TGTTCACGCTTACAAAGAAATGTTGACATATAATGCTCTGCTGTTTCATGGAGAG 422

720 TGT TTT CTC GTT CTT GAG AAT GCG TCG CAA AAT TTA CCA TTA TTG GTT CAC GGA GAG 779

423 GTTACTAATCCTGAGGTTGACATGTTGATAGAGAAAAGTATTCAATTGAACGGTTCTA 482

780 GTTACAGATTCAACGTTGATATTTTTCACCGAGAAAAAGCTTTATTGAAACAATTTTA 839

483 AGACCGTGGTGAGAAATTCACAAATTGAAGTCGTGATGAGCATTGTACCACATT 542

840 ATGCCITTAATTACAAGACTTCACAGCTCACACTGGTCAGC-GCCAACCACTACTGCA 898

543 GATGCTGTTAGTTTGTGATCTTTCACCTGAAGATTGTTGACGCAACTGTCAACCCCA 602

899 GATCGTTAATTTGAGACTCTTCCAAGAGGTTATGTAGACGCACTCTTACCA 958

603 CAACATCTTGTTTTGAACAGGAATTCCTCTTCCAAAGGGGCTTACAACCGGTAATTAA CCG

959 CAGCATCTTCTTCGATCGTAGTCTTGCTGTCCAAAGCTGGCCTTAACCCGCAATTTCCTTTCT
062

663 TGCCCTTCAGTCTCAAAAAGAAGATTCACAAACCACCCGCGGGCATTGTTGT

[illegible]

Db 1019 TGTCTTCAGTGTCTCAAAAGAGAGATCCATAGACAGGCTATTGTTGGCTGTCACTAGT 1078
Qy 723 GGAAGTAAAGATTTTTTCTTGGAGCTGATAGTCTCTCATGTATAGACGAAGAAAAG 782
Db 1079 GGAAGTAAAGATTTTTTCTTGGAGCTGATAGTCTCCACATGATAGGCGTAAAGAAA 1138
Qy 783 TGTCTTGTGTGATGTGTGTGATTTTACATGACACTGTAGCTTGTGTGATATATGCAAG 842
Db 1139 TGTTCCTGTGTGATGTGTGTGATTTACATCACTCAGCGGTGTCTATCATATATGCAAA 1198
Qy 843 GTGTTGAAAAGAAAAGACGACGACAGCTTGAACATTTCACTAGCTTCAATGACCA 902
Db 1199 GTTTTAAAGAGCTGTGTGACATTTAGCTAGAGCTTTTACAGCTTTAAGGACCT 1258
Qy 903 GATTTTATGGCTTCTTACAGACACTCAAAAGATTAAGTTAGTAAAGCGCCATGAG 962
Db 1259 GACTTCTATGGCTTCTTACAGACACTCAAAAGATTAAGTTAAGTAAAGCTTGTGAAA 1318
Qy 963 GTACCCGAATCTTTTCTTATGATCAAGAGATTTATCCAGTTTGTGTGTGAATG 1022
Db 1319 GTACTGATTAATTTGTATTTTCAATTTGAGACATCGTTCCATGTTTGTGTGAAACC 1378
Qy 1023 CTGACTGTGTGCGGCTCTCTCGA 1049
Db 1379 CTGATATGAGAGCATTGCTTTTGA 1405

RESULT 3

US-10-437-963-86830
; Sequence 86830, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 86830
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_85834C.1
US-10-437-963-86830

Query Match 43.2%; Score 549.2; DB 18; Length 1545;
Best Local Similarity 70.4%; Pred. No. 4e-147;
Matches 734; Conservative 0; Mismatches 308; Indels 0; Gaps 0;

Qy 11 GGAGCTCATATCAACAACCGATGATGGCATCTTATCTCCGTAGTGGATGTTCT 70
Db 153 GGAGCTCAACATCAAGCGGCGCGAGCACTGGCACTTCCAGCGGAAGCGGATGTTCT 212
Qy 71 TAAAGCATTTCTCTCTCAAGTGCATCATCTTTGGAGGGGCAATAGTCAATCCAAATTT 130
Db 213 CGCGCGCGTGTCTCCCAAGCGGATGATTTTGGAGGGGCAATGATGCCAACCT 272
Qy 131 GAAGCTCTATCACTACCACTGCTGCTGTATAGATACCGGGAGCGGATTTGAATC 190
Db 273 GAAGCCACCGGTGACACGACGACGAGCTTTGAGATACAGGAGGAGATCTTCAGGAC 332
Qy 191 TTTACTGTGTATGATGATTTAAACCCCTTATGACATTTTATGACAGATACAACAG 250
Db 333 GCTGCGCGCTGGAGCACTTTGTGCGCTCATGACGCTTATCTCAAGGACAAATATG 392

Qy 251 TCCTATGAAATCAAACTAGCAAGAGAGACAGGTGATTTTGGGTGAAGTTGATCC 310
Db 393 CCCAGAGAGATCAAGCTCGCAAAAAGAGTGGCGTGTCTTGTGTGAAGTTGATCC 452
Qy 311 TGTGTGTGCGCAGCAAAATTTCTCAAGTGTGATGATCTTTTCCGGAAGTTTACC 370
Db 453 TTTGTGAGCAACTACCAACTCCCAAGATGTGTGATATATTTGGGAAGTGTGCC 512
Qy 371 AGTTTCAAGAAATGTTGAGCAATATATGCTCTGTGTGTATGAGAGTTACTAA 430
Db 513 TGTCTTGAAGAAATGCTGAGGAGAGATGCTTGTGTGTATGAGAAATATTTGGACACT 572
Qy 431 TCCTGAGCTGATATGTTATATGAGAAAGTATTCATTTGAACGGTTCTAGACGCT 490
Db 573 TCAGCATGTGTGATCCTTTATGTGTGAAGGTTTCATTTGAGAAATATTTGGACACT 632
Qy 491 GGTGCAAGAAATTTCCAAATTTGAAGGTGTGATGAGACATGTTACCACTATGATGCTGT 550
Db 633 TGTACAAAGACTGCGACAGCTGAAAATTTGTATGAAACATATACCACTATGATGCGGT 692
Qy 551 TAAATTTGTTGAATCTTGTGACCTGAAGATTTGTGACGAACTGTCAACCCCAACATCT 610
Db 693 TAACTTGTGAATCATGCAAAAGAGTCAATGTGTGTGCAACGCTGACTCCCAACATCT 752
Qy 611 TGTTTGAACAGAAATCTCTCTTCCAAAGGGGCTTCAACCCGATTAATTAAGCTTCC 670
Db 753 TCTTCTCAACAGAAATGCTTATTTTCAAGGTTGTGAGCTTGTGACACACATATTTCTTACC 812
Qy 671 AGTCCCTCAAAAGAGATTCACAGAGAGCACTGTGTCACTGTATCAAGTGAAGTAA 730
Db 813 AGTACTAAAAGAAACTCATGACAGCAATTTGTCTGTGCGTTAACTAGTGAAGTAA 872
Qy 731 AAGATTTTCTTGGAGCTGATAGTGTCTCTCATGATAGCAAGAAAGAGTGTCTTG 790
Db 873 ACAATACTTCTTGTGTGTGACTGACAGTGTCTCTCATGATTAACGGAAGAAATCTCTGT 932
Qy 791 TGTATGTGTGTATTTTACAATGCACTGTAGCTGTGTATATGCAAGAGTGTGTA 850
Db 933 TGTGTGTGCAAGAAATATATAGCGCTCTGTGTCTTCTTTATGGAAGATTTGCA 992
Qy 851 AAAGAAATGCACTGACCAAGCTTGAAGCATCTTCACTAGCTTCAATGAGACCGATTTT 910
Db 993 ACAGCGCGTGTCTTGTATAGCTAAGCAATTTACAAGCTTCAATGCGCCAGATTTT 1052
Qy 911 TGGCTTCTTGAAGCAACTCAAAAGATTAAGTGAAGAGCGCATGGAAGTACCGCA 970
Db 1053 TGGCTTCCAAAGCACTTCAAGATTTGTCTTGAAGAGCGCTGGAAGTACCTGCA 1112
Qy 971 ATCTTTTCTTATGATCAAGAGATATTTCCATGTTTCTGTGTGAAGTGTGACTG 1030
Db 1113 TACATACTCATATATGTTTCAAGAGATGCGCTATGTTTATGTGCAACACCTCGAATG 1172
Qy 1031 GTTGGCGGCTCCTCTGAGAA 1052
Db 1173 GCTTCCATCGATCAACTTGA 1194

RESULT 4

US-10-425-115-86204
; Sequence 86204, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

Query Match	41.8%;	Score 531.6;	DB 18;	Length 1893;
Best Local Similarity	69.6%;	Pred. No. 5.3e-142;		
Matches 720; Conservative	0.0	Mismatches 314		

11 GGAGCTCTAATCAACAACAACCTGATGATTTGGCATCTTCACTCCGGATAGTGAATGCT 70
Db 457 GCAACTGATCATCACTCCGCCCGACGACTGSCACTCCATCTTCCGGAGGCGGTCT 516
QY 71 TAAGCAGTGTGCTCTCAACAGTGCACATCACTTTGGAGGGCAATGATGCACAAATT 130
Db 517 CGAGGCTGTAGTCCACACACGCCGAGGCACTTTGGAGGGCCATCATATCCCAACT 576
QY 131 GAAGCTCTTATCACTACACTGCTGCTGTGCTGTAGCAATACGGAGGCGCATTTGAATC 190
Db 577 GAAGCCGCGGTGACCAACACGCGCGCGGCGGTGAGTACAGGAGGAGATATAGAGGC 636
QY 191 TTACTCTGTATGATGATTTCAACCTCTTATGACATTTATTTGACAGATACAACAG 250
Db 637 GCTGCCCGCAGGAGTACTTTCAGCCGCTCAATGACATATCTCAACGACACAGAG 696
QY 251 TCTATGGAAATCAAACTAGCAGAGAGAGCCAGTGCATTTGGGGTGAAGTTGATCC 310
Db 697 CCCAGAGGAGATCAAGCTCCGAGAAAGATGATTAATTTTGTGCTGAAGTTGTATCC 756
QY 311 TGCTGTGCGCCACGACAAATTTCTCAAGTGAAGTGAATCTTTTCGGAGATGTTTACC 370
Db 757 CGCTGGACCACTACCAATTTCCCAATATGATGCTCATGATATATTTGGGAAGTCTTGCC 816
QY 371 AGTTCTACAAAGAAATGTTGATGATTAATATGCTGCTGCTGCTCATGAGAGTTACTAA 430
Db 817 GGTCTCTGAGAGATGTTGTGAGGAGAAATGCCGTTGCTGTTCAATGAGAAAGTCACAGA 876
QY 431 TCCGAGAGTTGACATGTTTGTATGAGAAAGGATTTCAATGAAACGGTCTTAAGACCGTT 490
Db 877 TCCACATGTGACACATTTGACCGCGAGAAAGGTTTTCATGACAAATATTAGACACACT 936
QY 491 GGTGCAGAAATTTCCACAATGAAAGGTGTGATGAGAGATGTTACCACTGATGCTGT 550
Db 937 TGTACAAAACTCCGACACTGAAATTTGTCAATGGAACATATCAACATATGATCACT 996
QY 551 TAAGTTTGTGAATCTTGCACTGGAAGATTTGTTGACAGCAACTGTCAACCCCAACATCT 610
Db 997 GAATCTCATAGATCATGTGAAGAGTGCATGTTGCTGCAACATGACTCCCAAGATCT 1056
QY 611 TGTTTTGAACAGAAATTCCTCTCTCCAGAGGGGCTTCAACCCGATATATTAAGCTTCC 670
Db 1057 CCTCCTCATATGAGAAATGCTTTATTTCAAGGGGTGCTTGCACCCACACAAATTAATGCTTCC 1116
QY 671 AGTCTCAAAAGAGATCCAAGGAGGCACTTGTGCACTGATGATGAAGAGGAAAGTAA 730
Db 1117 AATACTGAAAAGAGAGCTCATATGACAAGATTTGTCTGCTGTAAACAAATGGAGTAG 1176
QY 731 AAGATTTTTCTTGGGAGTGAATAGTCTCTCATGATGAGCAGAAAGAAAGAGTGTCTTG 790
Db 1177 ACGGTACTTCTTGGTACTGACACGGCTCCCATGATTAACGGAACAAAGATGTTCTCG 1236
QY 791 TGGATGTGCTGTATTTTACATGACCTGTAGGCTGTGTCAAGTATATGCAAGGTGTTGA 850
Db 1237 TGGATGTGCTGTATATATAGCGTCTGTGCTGTGCTCTTTATATGGAAGATTTGA 1296
QY 851 AAGGAAAAATGCACTCGACAGCTTGAAGCATTTCACTAGCTCAATGACACGATTTTGA 910
Db 1297 AAGGCTGTGTCCTTGTATTAACAGAGATTTACAGCTTCAATGACCTGATTTTGA 1356
QY 911 TGGCTTCTTAGAACCACTCAAGATTAAGTTGAGTAAAGCCGCAATGAAGGTATCCGA 970

Accession	Sequence	Position
Db	1357 TGGCTTCCCAAGAACCTTCAAGATTATCTCAGAAAGATCCTGGAAAGTTCTCGC	1416
QY	971 ATCTTTTCTTATGCATCAGAGATATTATATCCATGTTGCTGTGTAATGTCGACTG	1030
Db	1417 TACTTATTAACAGACGTTCAAGGAGATTGTGCTATGTTTAATGGACGCCCTTGATG	1476
QY	1031 GTTACCGGCTCCTC	1044
Db	1477 GCTTCCATCTGATC	1490

```

RESULT 5
US-10-425-115-86207
/ Sequence 86207, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: Ia Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 86207
/ LENGTH: 1258
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577__178630C.1
US-10-425-115-86207

```

Query Match	Best Local Similarity	Score	DB	Length
Matches	483; Conservative	0; Mismatches	230; Indels	0; Gaps
QY	11 GGAGCTCTCAATCACACAACTGATGATTTGGCATCTTCATCTCCGTATGCGTATGTTCT	27.1%;	345;	1258;
Db	309 GCAACTGATCATCACTCTCCGCCGAGACTGGACCTTCATCTTGGCGAAGGCGGTGCTT	67.7%;	18;	70
QY	71 TTAGCAGTGTCTCTCAACAGTCACATCACTTCTGGAGGGCAATAGCATGCCAAATTT	Pred. 2.7e-88;		
Db	369 CGAGGTGTAGGCCACAACGCGGAGGCAATTTGGAGGGCCATTCATGCCCCAACT	No. 2.7e-88;		
QY	131 GAAAGCTCTTATCACTACACACTGCTGCTGTGTATGACATCCGGAGGCGATATTGAATTC	Score 345;	DB 18;	Length 1258;
Db	429 GAAAGCTCTTATCACTACACACTGCTGCTGTGTATGACATCCGGAGGCGATATTGAATTC			
QY	191 TTTACCTTTATATGTGATTTCAACCTCTTATGACATTTATTGACAGTACCAACG			
Db	489 GCTGCGGCGAAGGAGTGTGCTTGGAGGCGCTCATGACATTCCTCAACACACACAG			
QY	251 TCTATGGAAATCAACATGACAGAGAGAGACAGCTCGTATTTGGGGTGAAGTTGTAACC			
Db	549 CCGAGAGAGATCAAGCTCGAGAGAGAGAGTGTATATTTTGTGTGAAGTTGTAATCC			
QY	311 TGTGTGGCCACGACAAATTTCTCAAGTGTGAGTACATCTTTGGGGGAAGTTTAC			
Db	609 TGTGTGGCCACATCAATTTCCAAAGTGTGTCACTGATATATTTGGGAAGTGTGCC			
QY	371 AGTTCTACAGAAATGGTTGAGATATATGCTCTGTGCTTTCATGAGAGGTTACTAA			
Db	669 GGTCTGAGAGAGATGGTCAAGCAGAGAAATGCCGTGTGTTTCAATGAGAAAGTCAACA			
QY	431 TCTGAGGTTGACATGTTGATAGAGAAAGATTCATTTGAACCGTTCTTAAGACGTT			
Db	729 TCCACATGTTGACACATTTGACCGGAGAGGTTTTCATTTGACAAATATTTGACACACT			
QY	491 GGTCGGAATTTCCCAATTTGAAGTGTGTATGAGACATGTTACACCACTGATGCTGT			

Db 789 TGACAAAACTCCGAGCTGAAATTTGTCATGAAACATATTCACAACATATGATGACAT 848
Qy 551 TAAGTTGTTGAATCTTGGCACTGAGAGATTTGTTGACGAACATGTGACCCCAACATCT 610
Db 849 GAACCTCATGAAATCATGTGAAGAAGTATGTTGGCGAAGAGTATGCTCCAGCATCT 908
Qy 611 TGTTTGAACAGAAATTTCTCTTCCAGAGGGGCTTACAAACCGCATATATTCGCTTCC 670
Db 909 CCTCTCAATAGAAATGCTTATTTTCAGGGTGGCTTGCACCCACACATATTTGCTTGC 968
Qy 671 AGTCCCAAAAGAGATTCACAGGAGGACCTTGTGCTGATGTAACAAGT 723
Db 969 AGTACTGAAAAGAGACTATAGTATGTTGCCATGAGCTCTCTATAGAAATG 1021

RESULT 6
US-10-767-701-2040
; Sequence 2040, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767, 701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 2040
; LENGTH: 931
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS34448_1
US-10-767-701-2040

Query Match 24.5%; Score 311.8; DB 18; Length 911;
Best Local Similarity 69.0%; Pred. No. 8.4e-79;
Matches 427; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

Qy 11 GGAGCTCTCAATCAACAACCTGATGATGGCATCTTCTCCGTATGATGATGTTCT 70
Db 313 GCAACTGGTCACTACTCGGCCGACGATGGCATCTTCCATCTTGATGAAGCGGTGCT 372
Qy 71 TAAAGGAGTTGCTCTCAAGTGCACATCACTTTGGAGGGCAATAGATGCAAAATTT 130
Db 373 CGAGGCTGTGCTCCCGCGAGCGAGGCAATTTGGAGGGCATCATCATGCTCCAACTT 432
Qy 131 GAAGCTCTCTATCACTACACTGCTGCTGTAGCATACCGGAGGCGATATTTGAATC 190
Db 433 GAAGCCCCCGGTGACCAACAAGCGCGCGGTGGAATACAGGGAGAGATATGAGGCG 492
Qy 191 TTATCTGTGATAGTATTTCAACCTCTTATGACATTTATTTGACATACCAACG 250
Db 493 GCTGCGCGCAGAGTACTCTGAGCCGCTCATGACATGTAACCGACAGCAACACAG 552
Qy 251 TCCATATGAATCAAACTAGAGAGAGAGAGCCAGTCTTATTTGGGGTGAAGTTGATCC 310
Db 553 TCCAGAGAGATCAAGCTCGCAAGAAAGATGTATTAATCTTGCTGTGAAGCTATATCC 612
Qy 311 TGCATGTCACAGCAAAATTTCTCAAGATGAGTGAATCTTTTCGGGAAGTTTATCC 370
Db 613 TGCATGAGCAATCAATTTCCCAAGATGGGTCTATATATTTGGGAAGTGTGCTCC 672
Qy 371 AGTTTACAAAGATGTTGAGCATATATGCTCTGCTGTTCTATGAGAGGTTTACTAA 430
Db 673 AGTCTCCAGAGATGATGATGAGCAAGAAATGCAATGCTTGTTCATGAGAGTACAGA 732
Qy 431 TCCGAGGTTGACATGTTGATAGAGAAAGTATTCATTTGAAGAGGTTTAAAGCCGTT 490
Db 733 TCCATATGTTGACATCTTGTGACCGTGAAGAGGTTTTCATTTGACAAATATTTGGCACCT 792

Qy 491 GGTGCAAAATTTCCAAATGAAAGTCTGTATGAGCATGTTATACCACTGATGATCTGT 550
Db 793 TGACAAAACTTCCAGAGTGAATTTGTCTATGAAACATATCAACATGATGATGAT 852
Qy 551 TAAGTTGTTGAATCTTGGCACTGAGAGATTTGTTGACGAACATGTGACCCCAACATCT 610
Db 853 GAACCTCATGAAATCATGTGAAGAAGTATGTTGGCGAAGAGTATGCTCCAGCATCT 912
Qy 611 TGTTTGAACAGAAATTTCT 629
Db 913 CCTCTCAATAGAAATGCT 931

RESULT 7
US-10-282-122A-41490
; Sequence 41490, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangshu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41490
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Vibrio cholerae
US-10-282-122A-41490

Query Match 22.1%; Score 281; DB 17; Length 1053;
Best Local Similarity 56.8%; Pred. No. 6.9e-70;
Matches 578; Conservative 0; Mismatches 430; Indels 9; Gaps 3;

Qy 15 CTTCATACACAACTGATGATTTGGCATTTATCTCCGTATGATGATGATGATGATGAT 74
Db 34 CTGACGATATACACCCCGAGAGACTGGACATTTGCGCATGGAGATGATGATGATGATG 93
Qy 75 GCAATGCTCTCAAGTGCACATCATTTGGAGGGCAATAGCATGCCAAATTTGAAG 134

439 GCGGTGACATCTTTATGATGTGAAGCCGTTTATTGACCAATTTTGAAGCCATTGCG 498

495 CAGAAATTTCCAACTTGAAGGTCGTGATGAGCATGTTACACCATGTATGCTTTAG 554

499 CAAAATTTCCCAACTAAAATTTGCTTTGAGCATATCAGACCAAGATGGGGAGAT 558

555 TTTGTGAATCTTGCACTGAAGATTGTGAGCAACTGTACCCCAACAATCTTGT 614

559 TATGTGTGCAAGCA--ATCGTTTCCTTGGGGCAACGTCACGCCACAATCTGATG 615

615 TTGAACAGAAATTCCTCTTTCAGAGGGGCTTACACCGCATATTACTGCTTCCAGTC 674

616 TTTAACCGCATCAATGCTGTGAGCGGTATTCGCCCCCACTTGTCTGCTCCATA 675

675 CTCAAAAGAGATCCACAGAGGAGCACTTGTGATGATTAACAAGTGAATTAAGA 734

676 TTGAAGCGACACCCATCAGCAAGATTGGCGGAGCGTCGACAGTGTCTGATGCG 735

735 TTTTCTTGGAGCTGATAGTGCTCTCATGATTAACGAAAGAAAGATGTTCTTGGA 794

736 TTTCTCTTGGACCGATTCAGTCTCCCATGCGAAATCGTAAGATCATCTTGGCGC 795

795 TGTGCTGTTATTTACATGACACCTGTAGCCTTGTCAGTATATGCAAGATGTTGAAAG 854

796 TGTGCGGGTATTTCAACGCCCAAGGGGATTTGCCGTTATGCTTCGCTGTTGAGGA 855

855 GAAATGCACTCGACAAAGCTTGAGACATTCATAGTCTCAATGAGACCGATTTTATGG 914

856 CTGAATCATTTGCAACATCTGAAAGGTTTGGGCTTTAAAGGCCCAAGATTTATGGC 915

915 CTTCCTAGAACAACTCAAAAGATTTAAGTTGAGTAAGAAGCCATGGAAGTACCGGAATCC 974

916 TTGCTGTATATGATGACGTGTGTGAATGGTTGCACTTCATTCCTTGAGCAGAAAG 975

975 TTTTCTTATGATGAGAGATATTATTCATGTTTGGCTGGGAATGCTGCATGCT 1032

976 ATCCCATTTAGCAATGATGCGTATATTCCTTTCTTGGGGTCAAAAGCTTATTTGGT 1033

```

RESULT 9
US-10-282-122A-8516
/ Sequence 8516, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haeselbeck, Robert
/ APPLICANT: Ohlsen, Karl
/ APPLICANT: Zysekind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23

```

	Prior Application Number: 60/253,625	
	Prior Filing Date: 2000-11-27	
	Prior Application Number: 60/257,931	
	Prior Filing Date: 2000-12-22	
	Prior Application Number: 60/267,636	
	Prior Filing Date: 2001-02-09	
	Prior Application Number: 60/269,308	
	Prior Filing Date: 2001-02-16	
	Remaining Prior Application data removed - See File Wrapper or PALM.	
	Number Of SEQ ID NOS: 78614	
	SOFTWARE: PatentIn version 3.1	
	SEQ ID NO 8516	
	LENGTH: 1032	
	TYPE: DNA	
	ORGANISM: Acinetobacter baumannii	
	US-10-282-122A-8516	
Query Match	21.0%; Score 267.4; DB 17; Length 1032;	
Best Local Similarity	56.4%; Pred. No. 5.7e-66;	
Matches	540; Conservative 0; Mismatches 411; Indels 6; Gaps 2	
Dn	27 CAACCTGATGATTGGCATCTTCATCTCCGTGATGGTGAATGTTCTTAAGGCAGTGTCTCT	86
Dn	22 CAGCCAGATGATGGCATGCACATTTAGTGATGGTTTAGATTAACGTACTCTTCCA	81
Dn	87 CACAGTGCATCACTTTGGGAGGCGAAATGATGCCAAATTTGAACCTCCTACT	146
Dn	82 GATTTGGTTAACATTTGCCCGTCGATTTGTATGCTTACCTGTACCCGCTGTAAA	141
Dn	147 ACCATGCTGCTGTGTGATGATACC GGAGGCGATATTTGAATCTTTA CTGTTATGT	206
Dn	142 ACAGTAGAAGAAGCTTTAGCTTATGCGAAGCATTTCTTGTCATGTTCCGAAAGCAT	201
Dn	207 GATTTC AACCTCTTATGACATTTATTTGACAGATACACAGTCCATGGAATTC AAA	266
Dn	202 AATTTTGACCCCTCGATGATGCTTTATTTATTTATGTGACCACTTCCACGATGAAGTTGCT	261
Dn	267 CTAGCAAGAGAGCGCAGGTCGTA TTTGGGGTGAAGTTGATCCCTGCTGTCGCCAGCA	326
Dn	262 AAATATTAAGATTCAGAACATGTAATAATGCGATTAACTTATCTGCTGTCGACACACA	321
Dn	327 AATTTCAAATGAGTGAAGTACTGATCTTTTCGGAAAGTTTAA CAGTTCTACAGAAATG	386
Dn	322 AACTCGGTAATGATGATGAGTATTT -- CGTAAGTTTATGACAGTCA TTGAGCAATTA	378
Dn	387 GTTGAGCATATATATGCTCTGCTGTTCTATGSAAGAGTTACTATATCCATGAGTTGACATG	446
Dn	379 GAAGGAGCATCAAGTTCCTGTTATGCTTCAATGTTGAAGTAACTCACATATCATGATGATAT	438
Dn	447 TTGTATGAGAAAGATATCTCATTTGAACCGTTCTPAAGCCGTTGGTGCAGAAATTTCCA	506
Dn	439 TTGTATCGTAAAAACGCTTCTTAATGAGGATATATACCTCTATTTGAACAGTTCCCG	498
Dn	507 CAATTGAAGTGTGATGAGCATGTTTACACACATTTGATGCTGTTAAAGTTTGTGATCT	566
Dn	499 AAACCTTAAGTTGTGCTTGAAGCATCACTACACAGATGACAGACACTTGT--TTA	555
Dn	567 TGCACTGAAGATTTGTTGTCAGCAACTGTACCCCCACACATCTTGTTTTGAACAGGAT	626
Dn	556 GAACAAGACCGTAATGTGGCGGCACAACTACTCCACCAATTTATTTAACCGTAAT	615
Dn	627 TCCTCTTCCAGAGGGGGCTTACACCGCATTAATTAATGCTTCCAGTCCCTCAAAAGAG	686
Dn	616 GATATGTTGTTGGTGTGATTTAAACCGCATTTTACTGTTTACCAATTTTAAAGCTCAA	675
Dn	687 ATCCAAGAGGAGCATTTGTCACTGTACGCTGTAACTGAGTGAAGTAAAGATTTTTTCTTGGG	746
Dn	676 ACACATCAAAACAATTTGCTTGAAGTGGCAGCAAGGCGTAAATCTTAAATTTTCTTAGGT	735
Dn	747 ACTGATATGCTTCCATCATATGACGAAGAAAAGATGTTCTTGTGATGTGCTGATAT	806
Dn	736 ACGAAGAGGCTCTCATGACAAATGCAAAAGAGATGCTTGTGATGTGACAGCTGC	795

QY 807 TACATGACCTGTAGCTTGTGATATGCGAAGGCTTTGAAAAGAAATGCACTC 866
DB 796 TATAGTGACCAATATGATTAAGCTTTATGACAAAGCATTTGACAGGTGCGTAATTA 855
QY 867 GAAACCTTGAGACATTCATCTGATTCATAGACAGATTTTATGAGCTTCCCTAGAAC 926
DB 856 GACCGTCTGAAAGTTTGTCTAGCATTTTGGTGGAGCTTACGGCTTACACAGTAAT 915
QY 927 AACTCAAGATTAACTGATGAGACGCGCATGGAAGGTACCGGATCTTTCTTAT 983
DB 916 ACTTTCACCATTACTTTGGTTAAAGAAATACCTCGTTCAGAACTCTTTGATTAAT 972

RESULT 10
US-10-282-122A-14242
; Sequence 14242, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14242
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Burkholderia mallei
US-10-282-122A-14242

Query Match 20.9%; Score 265.2; DB 17; Length 1056;
Best Local Similarity 54.4%; Pred. No. 2.5e-65;
Matches 559; Conservative 0; Mismatches 463; Indels 6; Gaps 1;

QY 14 GCTCTCAATCAGCAACCTGATGATTTGATCTCGTATGATGATTTCTTAA 73
DB 24 GCTGACCTCTGCGCGCGCGGACCTGCACTGCACTGCGCAGCGCGGATGCTCGC 83
QY 74 GCGAGTGTCTCTCAGTGCATCATCTTTGGAGGCGCAATATGTCACCAATTTGAA 133
|||||

DB 84 GCGCTCTCTGCCGACACCGCGCTCGCAGTTTCGCGCGCGCATCATGCGCAACTGAA 143
QY 134 GCTCTCTACATCACTCACTGCTGCTGTGATGATATACCGGAGCGATATTTGAATCTTT 193
DB 144 GCGCGCGGTCAACGACGACCGCGGCGAGGCGCTTACCGGACGATCTGCGCGCGGT 203
QY 194 AACTTTGATATGATTTTCAACCTCTTATGACATTTTATGACAGATTAACAAGATCC 253
DB 204 GCGCGCGGACGATGACGCTTTCAGCGCTTATGACGCTGATCGTACCGAACAACCGCGC 263
QY 254 TATGAAATCAAACTAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 313
DB 264 CGACGAATATCGCGCGCGGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 323
QY 314 TGGGCGACAGCAAAATTTCTCAAGATGAGATGATCTTTTGGGAGAGTTTACAGT 373
DB 324 GGGGCGACAGCAAACTCGGACGCGCGCGTACCGACCTGCTCGGAGAGTCCGAGAGAC 383
QY 374 TCTACAGAAATGATGATGATATATGCTCTGCTGCTGATGATGAGAGAGTTACTATCC 433
DB 384 GCTCGAGGAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCC 443
QY 434 TGAAGTTGACATGTTTATGAGAGAGAGATTTGAAACGGTTCTTAAAGCCGTTGT 493
DB 444 GTGATGACCTGTTTGAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 503
QY 494 GCAGAAATTTCCAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 553
DB 504 CCGCGCGCTCGCGCGGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 563
QY 554 GTTTGT-----TGAATTTGACTGAGAGATTTGTCAGCAACTGTCAACCCGCAACA 607
DB 564 CTACGTGCGCATCGACGCGCGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 623
QY 608 TCTGTTTGAACAGAAATTTCTCTTCCAGAGGCGCTTACAGCGCATTAATTAATGCT 667
DB 624 TCTGCTGACAAACCGCAATCGATGTTTTCGCGGAGATCCGTCACATTAATGCTGCT 683
QY 668 TCCAGTCTCTCAAAAGAGATTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 727
DB 684 GCGGTCTCTAAGGCGAG 743
QY 728 TAAAGATTTTCTTGGGACGTATGCTCTCTCTATGATTAAGAGAGAGAGAGAGAG 787
DB 744 TCGCGCTTCTCTCTCGGACGACGACGCGCGGACGAGAGAGAGAGAGAGAGAGAG 803
QY 788 TTGATGATGCTGTGATTTTAAATGACCTGTAGCTTGTCAATATGAGAGAGAGAG 847
DB 804 GTGCGGTGCGCGGCTGTACACCGCGCTGACGCGCTGACGCTGTACGCGAGAGAGAT 863
QY 848 TGAAG 907
DB 864 CGACAGAGGCGGCGCTGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923
QY 908 TATAGGCTTCTTGAAGCACTCAAGATTAAGTTAGTTAGAGAGAGAGAGAGAGAG 967
DB 924 CTAGGTTTTCGCGCAGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983
QY 968 CGAATCTTTTCTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1027
DB 984 GCGGAGATGACGCGGCG 1043
QY 1028 CTGCTTCG 1035
DB 1044 CTGCGCGC 1051

RESULT 11
US-10-282-122A-11985
; Sequence 11985, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Hasebeck, Robert
APPLICANT: Ohlssen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Foreyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282.122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11985
LENGTH: 1092
TYPE: DNA
ORGANISM: Burkholderia cepacia
US-10-282-122A-11985

Query Match 20.8%; Score 264.4; DB 17; Length 1092;
Best Local Similarity 56.2%; Pred. No. 4,3e-65;
Matches 540; Conservative 0; Mismatches 411; Indels 9; Gaps 2;

14 GCTCTCAATCAACAACTGATGATGGCATCTTCATCTCGTGATGATGTTCTTAA 73
63 GCTTTCGCTCGCCGTCGCGACGACTGCACTGCGAGCGCGCATGCTGC 122
74 GGCAGTTGCTCTCAACATGCACTTTGGGAGGGGCAATGAGCCAAATTTGAA 133
123 GCGGTGCTGCGACACCGCCGCAATGCGCCGCCCATCGTCAATGCGCACTGAA 182
134 GCTCTCAATCAACAACTGATGATGGCATCTTCATCTCGTGATGATGTTCTTAA 193
183 GCGGTGCTGCGACACCGCCGCAATGCGCCGCCCATCGTCAATGCGCACTGAA 242
194 ACCTGTGATGATGATTTCAACCTCTTATGACATTTATTTGACAGATTAACAAGTCC 253
243 GCGGTGCTGCGACACCGCCGCAATGCGCCGCCCATCGTCAATGCGCACTGAA 302
254 TATGAAATCAAACTGATGATGGCATCTTCATCTCGTGATGATGTTCTTAA 313
303 GCGGTGCTGCGACACCGCCGCAATGCGCCGCCCATCGTCAATGCGCACTGAA 362
314 TGTGTCCAGCAAAATTTCTCAAGATGATGATGATGATGATGATGATGATGATGAT 373
363 AGGCGCCAGCAAAATTTCTCAAGATGATGATGATGATGATGATGATGATGATGAT 419
374 TCTACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433

420 GCTCGAGCGGATGACAGAAACCGCATGCCGCTGCTGTCAGACGCGAGGTGACCGATGC 479
434 TGAAGTTGACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 493
480 GTTCATGACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
494 GCGAAATTTCCAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553
540 CCGGATTTCCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
554 GTTGT-----TGAATCTTTCACCTGAAGATTTGTTGACAACTGTCAACCCCAACA 607
600 CTACGTCGCGGATGCGGAGCGGCGCGCGCGCTGCTGCGGCGGACATCAACCGGATCA 659
608 TCTTGTGTTGACAGAAATTTCTTCTTCAAGGGGCTTCAACCCGATTAATTAAGTCT 667
660 CCGTGTGATCAACCGGAAACCGGCTGTTGCTGCGGAGATCCGCGCATTAATTAAGTCT 719
668 TCGAGTCTTCAAAAGAGATTCACAGGAGGACCTGTGTCACTGTATCAAGTGAAG 727
720 GCGGTGCTGAGCGGAGCGGATGCGTGTGCGCTGTGTGAGGCGCGGACGCGGCA 779
728 TAAAGATTTTCTTGGACTGATGATGCTCTCATGATGACGAAAGAGTGTTC 787
780 CCGGCTTCTTCTGCGGACCGGACGCGGCGCGGCGGACGCGGAAACCGG 839
788 TTGTGATGCTGCTGATTTTACATGATGATGATGATGATGATGATGATGATGATGAT 847
840 GTGCGGTTGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
848 TGAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907
900 GCGACGCGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959
908 TTAATGCTTCTTGAAGCACTCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 967
960 CTACGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1019

RESULT 12
US-09-790-988-1
Sequence 1, Application US/09790988
Patent No. US20020127687A1
GENERAL INFORMATION:
APPLICANT: SHIGENOBU, SHUJI
APPLICANT: WATANABE, HIDEKI
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 20.8%; Score 264.2; DB 9; Length 640681;
Best Local Similarity 54.0%; Pred. No. 2.2e-63;
Matches 564; Conservative 0; Mismatches 478; Indels 3; Gaps 1;

20 AATCAACAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 79
368078 AATCAATTAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 368137
80 TGTCTCTCAAGTGCATCACTTTGGAGGCAATAGTCAATGCCAAATTTGAAGCTTCC 139

RESULT 13
US-10-282-122A-7036
; Sequence 7036, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu

```

/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Onlsen, Kari
/ APPLICANT: Zysek, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282.122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 7036
/ LENGTH: 1047
/ TYPE: DNA
/ ORGANISM: Escherichia coli
US-10-282-122A-7036

Query Match      19.3%; Score 245.6; DB 17; Length 1047;
Best Local Similarity 54.1%; Pred. No. 11e-59;
Matches 547; Conservative 0; Mismatches 459; Indels 6; Gaps 2;

QY      21 ATCACACAACGATGATGATGGATCTTCATCTCCGTGATGATGTTCTTAAGGACGTT 80
DB      28 ATCCGCCCGCCGAGACGACGTGCACCTTCACCTCCGCGATGCGACATGTTAAAACGTGTC 87
QY      81 GTCTCTCACAGGACATCATCTTTGGAGGGGCAATGTCATGCGCAATTGGAGCGCTCT 140
DB      88 GTGCATATACAGGAAATTTATAGACGGGCTATCGTAATGCCAATCTGGCTCCGCC 147
QY      141 ATCACTACACGCTGCTGCTGCTGATACATACCGGAGCGCATATGAAATCTTTACCTGTT 200
DB      148 GTGACACCGCTTGAGGCTGCGTGGCGTATCGCAGCGATTTCTTAGCGCGATACCTGCC 207
QY      201 GATAGTATTCACACCTCTTATAGACTTTATTTGACAGATCAACACAGTCTTATGAA 260
DB      208 GGGACGATTTACACCCATTGATGACCTGTATTTAAAGATTTGCTGATCTTAAATGAG 267
QY      261 ATCAAACTAGACAGAGAGACCGCAGTGTATTTTGGGAGTAAAGTTGACCTGTGGTGC 320
DB      268 CTGAGGCGCGGATTTTACGAAAGCGGTGTCAACCGTGCAAAACTTTACCGGCAACGCA 327
QY      321 ACGACAAATTCGAAGATGAGTGAATCTTTTGGGGAAGTGTATTCACAGTTCTTACA 380
DB      328 ACCACTAATTCGACCGACGCGTGAAGTCAATTGACGCA---TCATGCGGTACTTGA 384
QY      381 GAATGTTAGATATATATGCTCTGCTGTTATGAGAGGATTAATATCTGAGGTT 440

```

Db	385	CGATGGAAAAAATCCGTATGCCGTACTGGTGCATGGTGAAGTGCACATGCAGATATC	444
Qy	441	GACATGTTGTATGAGAAAAAGGTATTCAATTGAAACGGTTCTPAGACCGTGGTCAGAAA	500
Db	445	GACATTTTGTATGTTGATACCGCGCTTTATAGAAACCGTGAAGAACCTTCGCCACACGC	504
Qy	501	TTTTCCAAATGAAAGTCCGATGATGAGACATGTTCACACATTTAGTCTGTAAATTGTT	566
Db	505	CTGACTCGCGTGAATTCGTTTTTGAACACATCACACCAAGAAAGTCCGCCGCTATGTC	564
Qy	561	GAATCTTGCACTGAGAGATTGTTGCAGCACTGTACCCCAACAACATCTTGTTTGAC	620
Db	565	CGTACGGAATATAG---GCTGGCTGCACACATCACTCCGCAACATCTGATGTTTAC	621
Qy	621	AGGAATTCCTCTTCCAGGGGGCTTACAACCGCATTAATTAATGCTGCCCTCAGTCCATAA	688
Db	622	CGCAACCATATGCTGGTTGAGAGCGTGGCTCCGCACCTGTATGTCTTACCATCTCTAA	681
Qy	681	AGAGAGTCCACAGGAGGCACTGTGTCACTGTAAACAAGTGAAGTAAAGATTTTTT	740
Db	682	CGTAATATTCACCAACAGGCATTCGCTGAACCTGTCGCCAGGGGTTTTATACGATATTC	741
Qy	741	CTTGGGACGTATAGTGTCTCTCATATGATTAACGAAGAAAGATGTTCTTGTGATGATGCT	800
Db	742	CTCGATACGATTTCTGGCCACATGCACTCATCCGAAAGAGACAGTTGGCGCTGCCG	801
Qy	801	GGTATTTTCAATGCACCTGTAGCTTGTCAATATATGCGAGGTGTTGAAAAGAAAAT	866
Db	802	GGCTGTTTCAACCCCAACCGCGCTGGCGAGTTACGTCACCGTTTGGAAAGAAATGAT	861
Qy	861	GCACCTGCACAAAGCTTGAAGCATTCACCTACCTTCATAGACGAGATTTTATGGGCTTCT	920
Db	862	GCTTTGACGACCTTTGAAGCATTTCTGTTCTGTAAAGCGCCGCAAGTTCTATGGGTGCG	922
Qy	921	AGGAACAACCTCAAGATTAAGTTGAGTAAGCGCCATGGAAGTACCCGAATCCTTTTCT	980
Db	922	GTCAAACACATTCATCGAACTGTATGATGTAAGAAAGCAACAGTGTCTGAAAGCATCGCA	981
Qy	981	TATGCAATCAGAGATATTAATCCATGTTTGTGCTGTGAATCTCGACGTGGT	1032
Db	982	CTGACTGATATCACGCTGTGGCAATTCCTGCGCGGGAAACGTAAGCTGGT	1033

RESULT 14

```

US-10-282-122A-30305
; Sequence 30305, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

```

```

? PRIOR APPLICATION NUMBER: 60/230,347
? PRIOR FILING DATE: 2000-09-09
? PRIOR APPLICATION NUMBER: 60/242,578
? PRIOR FILING DATE: 2000-10-23
? PRIOR APPLICATION NUMBER: 60/253,625
? PRIOR FILING DATE: 2000-11-27
? PRIOR APPLICATION NUMBER: 60/257,931
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: 60/267,636
? PRIOR FILING DATE: 2001-02-09
? PRIOR APPLICATION NUMBER: 60/269,308
? PRIOR FILING DATE: 2001-02-16
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 78614
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 30305
? LENGTH: 1047
? TYPE: DNA
? ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-30305

```

Query Match	19.0%;	Score 241;	DB 17;	Length 1047;
Best Local Similarity	53.7%;	Pred. No. 2.4e-58;		
Matches 546;	Conservative 0;	Mismatches 465;	Indels 6;	Gaps 2;

OY	15	TTCTCAATCAACAACCGATGATTGGATCTTCATCTCCGTGATGGGATGTTCTTAAG	74
Db	13	CTCAACCTCTGCGCCCGGACGACTGGCATTCACTCGGTGAAGGCGCGCTGGCC	72
OY	75	GCAGTGTCTCTCAAGTGCACATCACTTTGGAGGGCAATAGTCATGCCAAATTGAAG	134
Db	73	AATACCGTCGGCAGCGCGCCGACCTTCGGCGCGCATCGTCATCCCAACTGTGTG	132
OY	135	CCTCCTATCACTACCACCTGCTGCTGCTGTACATACCGGGAGGGGATATTGAATCTTTA	194
Db	133	CCGCGCGTGGCCACGCGCGGAGCCGACGCTACCGGCAGCGCATCTCGCGCCGCT	192
OY	195	CCGTGTGATGATGATTTTCAACCTCTTATGACATTTATTTGACAGATACAACAGTCT	254
Db	193	CCCGCGCGCACCGCTTGAGCGCGCTGATGGTGTACTCAACGACCGACAGCAC	252
OY	255	ATGGAATCAACCTAGCAAGAGAGACGAGTCTGATTTTGGGTGAATTTGACCTCT	314
Db	253	GAGGAATCCGTCACCGCCAGGCCACGCGCTTCGTGATGCGCCGACCTCTATCGGGA	312
OY	315	GGTGCACGACAAATTTCTCAAGATGGAATGATCTGATCTTTTCGGAAAGTTTACAGTT	374
Db	313	GGCGCACACCACTCGGACTCTGGCGTGAAC--CGGCATGACAAATATCTTGAAGGG	369
OY	375	CTACAAAGATGGTTGAGCATATATATGCTCTGCTGTGTTTATGAGAGTTACTAATCT	434
Db	370	CTCGAGGGAGATGGCCGAGGTGGCGATGCGCGTCTGGTGGCAATGGGAGGTGACCGCGCC	429
OY	435	GAGGTGACATGTTGATAGAGAAAGGATTCATTGAAACGGATCTTAAGACCTTGGTG	494
Db	430	GAGGTGACATGTTGATGATCGGAGAAAGCATTCATGACAGACACTGCGCGGGTGGTC	489
OY	495	CAGAAATTTCCACATTTGAGGTCTGTATGAGCATGTTACCAACCATTTGATCTGTAA	554
Db	490	GAGCGCTTCCGACCTGAAAGTGTCTTTCGAGCATACACACCGGCGACCGCCCGCAG	549
OY	555	TTTGTGTGAATCTTGACATGAAAGATTTTGTGAGAGACATGACACCCCAACATCTTGT	614
Db	550	TTTCGTCCGGGAAGCCCC--GGCCAAAGTCGGCGGACCATTTACCGCCCATACCTTGCT	606
OY	615	TTGAACAGGAATTTCTCTCTTCAGAGGGGCTTACACACCGCATATTAATCTGCTTCAATC	674
Db	607	TACAAACCGAACCAATGCTGTGGTGGCGGTATCGTCCGCACTTCTATGCTCGCCGATC	666
OY	675	CTCAAAAGAGATTCACAGGAGGACATTGTGTACAGCTGTATCAAGTGGAAAGTAAAGA	734
Db	667	CTCAAGCGCAACCCACAGGAACCTCTGTGAGCGCGCGGTAGCGGCATCCGAA	726

Mon Feb 14 10:10:56 2005

us-10-070-277-1.rnpb

Page 13

Search completed: February 12, 2005, 07:01:22
Job time : 817 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 01:51:56 ; Search time 4610 Seconds
(without alignments)
10494.509 Million cell updates/sec

Title: US-10-070-277-1

Perfect score: 1271

Sequence: 1 tgcgaataatgagctctca.....tgaagattcactgataaaaa 1271

Scoring table: IDENTITY NUC

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hlc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_ges1.*
9: gb_ges2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	816.8	64.3	820	5	B0516583
2	767.4	60.4	1007	7	CK252266
3	744.6	58.6	751	5	B0516584
4	690.8	54.4	915	7	CK253198
5	689.6	54.3	790	7	CK242513
6	662.6	52.1	920	7	CK250612
7	658.2	51.8	725	7	CK249779
8	616.2	48.5	723	7	CK242510
9	607.6	47.8	708	7	CK251510
10	599.4	47.2	732	4	B1927946
11	595	46.8	957	7	CK242511
12	585.4	45.9	1301	3	CNS0A31G
13	583.8	45.1	1294	3	CNS0A38S
14	574.4	44.2	1297	3	CNS0A2XU
15	561	44.1	792	7	CK251764
16	555.2	43.7	1300	3	CNS0A3CO
17	550.8	43.3	1080	3	CL959705
18	549.6	43.2	583	1	A1895210
19	537.4	42.3	564	2	BE451410
20	514.4	40.5	532	2	BE451489
21	506.2	39.8	519	7	CK253199
22	500	39.3	954	7	CK247301
23	498	39.2	1296	3	CK248473
24					BX828477

25	496	39.0	905	7	CK246423	CK246423	EST730060
26	481	37.8	821	7	CK242512	CK242512	EST726149
27	465	36.6	957	7	CK245528	CK245528	EST729165
28	462.8	36.4	897	7	CK249778	CK249778	EST733415
29	441.4	34.7	1145	7	CK172140	CK172140	FGAS04769
30	441.2	34.7	641	4	B1203471	B1203471	EST521511
31	438.2	34.5	1153	7	CK170617	CK170617	FGAS04556
32	423	33.3	834	7	CK245538	CK245538	EST729175
33	421.4	33.2	1137	7	CK171201	CK171201	FGAS04635
34	404.2	31.8	743	5	B0636341	B0636341	049G06
35	403.2	31.7	887	7	CV130150	CV130150	MGfrc088
36	398.8	31.4	410	7	CK245539	CK245539	EST729176
37	370.2	29.1	1179	7	CK171534	CK171534	FGAS04683
38	365.6	28.8	1156	7	CK169818	CK169818	FGAS04446
39	362.2	28.5	1056	7	AY109986	AY109986	268
40	360.8	28.4	1174	7	CK171933	CK171933	FGAS04741
41	353	27.8	627	7	CF445164	CF445164	EST681509
42	346.8	27.3	607	2	BE344806	BE344806	946029B05
43	344.6	27.1	758	7	CV084527	CV084527	MGfrc1085
44	336.8	26.5	1093	7	CK171050	CK171050	FGAS04615
45	329	25.9	786	4	B0576697	B0576697	B0576697

ALIGNMENTS

RESULT 1
B0516583
LOCUS
DEFINITION
B0516583
EST623998 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STM1Y71
5' end, mRNA sequence.

ACCESSION
B0516583
VERSION
B0516583.1 GI:21375452
KEYWORDS
EST.
SOURCE
Solanum tuberosum (potato)
ORGANISM
Solanum tuberosum

REFERENCE
Bukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum.

1 (bases 1 to 820)
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J., and Karanaycheva, S.A.

Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
Other ESTs: EST623999

COMMENT
Contact: Robin Buell
The Institute for Genomic Research

7712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@ligr.org
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.

FEATURES

source

Location/Qualifiers
1..820
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultiivar="Kennebec or Binjle"
/db_xref="taxon:413"
/clone="STM1Y71"
/issue_type="mixed tissues"
/lab_host="SOLR"
/clone_id="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Combination of untreated and phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."

ORIGIN

Query Match 64.3%; Score 816.8; DB 5; Length 820;
 Best Local Similarity 99.8%; Pred. No. 1.1e-214;
 Matches 818; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 320 CACGACAAATTCCTCAAGATGAGTGAATCTTTTCGGGAAAGTGTTCACAGTTCTACA 379
 Db 1 CACGACAAATTCCTCAAGATGAGTGAATCTTTTCGGGAAAGTGTTCACAGTTCTACA 60
 QY 380 AGAATGTTGAGCATATATGCTGCTGCTGCTTCAATGAGAGTTAACTAATCTGAGGT 439
 Db 61 AGAATGTTGAGCATATATGCTGCTGCTGCTTCAATGAGAGTTAACTAATCTGAGGT 120
 QY 440 TGACATCTTTGATGAGAAAAAGGTATTCATTTGAAAGCGTTTCAAGACGTTGTCAGAA 499
 Db 121 TGACATCTTTGATGAGAAAAAGGTATTCATTTGAAAGCGTTTCAAGACGTTGTCAGAA 180
 QY 500 ATTTCACAAATTAAGAGTGTGATGAGCATGTTACACCAATTAATGCTTTAAAGTTGT 559
 Db 181 ATTTCACAAATTAAGAGTGTGATGAGCATGTTACACCAATTAATGCTTTAAAGTTGT 240
 QY 560 TGAATCTTGCACTGAAGATTGTTGACAGCACTGTCACCCCAACAATCTTTGTTGAA 619
 Db 241 TGAATCTTGCACTGAAGATTGTTGACAGCACTGTCACCCCAACAATCTTTGTTGAA 300
 QY 620 CAGGAATCTCTCTCCCAAGGGGCTTACACCCGATATTAATGCTTCCAGTCTCAAA 679
 Db 301 CAGGAATCTCTCTCCCAAGGGGCTTACACCCGATATTAATGCTTCCAGTCTCAAA 360
 QY 680 AAGAGATCCACAGGAGGCACTGTGTGACAGTGTAAACAGTGGAAAGTAAAGATTTT 739
 Db 361 AAGAGATCCACAGGAGGCGCTTGTGTGACAGTGTAAACAGTGGAAAGTAAAGATTTT 420
 QY 740 TCTTGGAAGTATGCTCTCTCATGATGAGCAAGAAAGAGTCTTGTGATGTGC 799
 Db 421 TCTTGGAAGTATGCTCTCTCATGATGAGCAAGAAAGAGTCTTGTGATGTGC 480
 QY 800 TGGATTTTACATGACCTGTAGCTGTGTGATGAGTATGAGGAGTGTGAAAGGAAAA 859
 Db 481 TGGATTTTACATGACCTGTAGCTGTGTGATGAGTATGAGGAGTGTGAAAGGAAAA 540
 QY 860 TGCATCGACAGCTTGAAGCATCTAGCTTCAATGAGCAAGATTTTATGGGCTTCC 919
 Db 541 TGCATCGACAGCTTGAAGCATCTAGCTTCAATGAGCAAGATTTTATGGGCTTCC 600
 QY 920 TAGGAACAATCAAGAATTAAGTGAAGAGCCATGAGAGTACCCGAATCTTTTC 979
 Db 601 TAGGAACAATCAAGAATTAAGTGAAGAGCCATGAGAGTACCCGAATCTTTTC 660
 QY 980 TTAAGCATCAGAGATATTAATCCCATGTTGCTGGTGAAGTCTGAGCTGGTCCGCG 1039
 Db 661 TTAAGCATCAGAGATATTAATCCCATGTTGCTGGTGAAGTCTGAGCTGGTCCGCG 720
 QY 1040 TCCCTCTGAGATCAATTTGTCATCTTGTGATGATTAATTTGATTCACCAAGATAT 1099
 Db 721 TCCCTCTGAGATCAATTTGTCATCTTGTGATGATTAATTTGATTCACCAAGATAT 780
 QY 1100 AGACTGTAGGTATCATCTTTTCTTCACTGTTGATTAGA 1139
 Db 781 AGACTGTAGGTATCATCTTTTCTTCACTGTTGATTAGA 820

RESULT 2
 CK252266 1007 bp mRNA linear EST 30-JUL-2004
 LOCUS ESTY35903 potato callus cDNA library, normalized and full-length
 DEFINITION Solanum tuberosum cDNA clone POCB81 5' end, mRNA sequence.
 ACCESSION CK252266
 VERSION CK252266.1 GI:39806092
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
 1 (bases 1 to 1007)
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 TITLE Generation of ESTs from potato callus tissue
 JOURNAL Unpublished (2003)
 COMMENT Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@igrr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via http://genome.arizona.edu/orders/
 Seq primer: ATG TAG GTG ACA CTA TNG.
 Location/Qualifiers
 1..1007
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /submitter="Kennebec"
 /db_xref="taxon:4113"
 /clone="POCB81"
 /tissue_type="callus"
 /lab_host="DH10B-Tona"
 /clone_lib="potato callus cDNA library, normalized and
 full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Solanum tuberosum var.
 Kennebec callus tissue grown on solid media."

FEATURES

source

ORIGIN

Query Match 60.4%; Score 767.4; DB 7; Length 1007;
 Best Local Similarity 99.2%; Pred. No. 5.4e-201;
 Matches 771; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 TGCAAAATGAGAGCTCTCAATACACACAACCTGATGATGGCATCTTCACTCCGATAG 61
 Db 221 TGCAAAATGAGAGCTCTCAATACACACAACCTGATGATGGCATCTTCACTCCGATAG 290
 QY 62 TGATGTTCTTAAGCAATGCTGCTCTCACTGACATCACTTTGGAGGCAATAGTCA 121
 Db 291 TGATGTTCTTAAGCAATGCTGCTCTCACTGACATCACTTTGGAGGCAATAGTCA 350
 QY 122 GCCAAATTTGAAGCTCTCACTACCACTGCTGCGCTGTAGCATATCCGGAGCCGAT 181
 Db 351 GCCAAATTTGAAGCTCTCACTACCACTGCTGCGCTGTAGCATATCCGGAGCCGAT 410
 QY 182 ATTGAATCTTTACCTGTTGATGATGATTTCAACCTCTTAATGACATTTATTTGACAGA 241
 Db 411 ATTGAATCTTTACCTGTTGATGATGATTTCAACCTCTTAATGACATTTATTTGACAGA 470
 QY 242 TACAACCAATCTATGAAATCAAACTAGCAGAGAGCCAGTCTGTTGGGTGA 301
 Db 471 TACAACCAATCTATGAAATCAAACTAGCAGAGAGCCAGTCTGTTGGGTGA 530
 QY 302 GTTGTACCTGCTGTGTCACAGCAAAATTTCTAAGATGAGTGAATCTTTTGGGGA 361
 Db 531 GTTGTACCTGCTGTGTCACAGCAAAATTTCTAAGATGAGTGAATCTTTTGGGGA 590
 QY 362 GTGTTACCAAGTTCTTACAGAAATGTTGAGATTAATGCTGCTGTGTTACAGAGA 421
 Db 591 GTGTTACCAAGTTCTTACAGAAATGTTGAGATTAATGCTGCTGTGTTACAGAGA 650
 QY 422 GGTACTAATCTGAGGTTGACATGTTGATGAGAAAAAGGTATTCATTTGAAAGGTTCT 481
 Db 651 GGTACTAATCTGAGGTTGACATGTTGATGAGAAAAAGGTATTCATTTGAAAGGTTCT 710
 QY 482 AAGACCGTTGGGCAAAATTTCCACAATTAAGAGTGTGATGAGAGATGTTACACCAT 541
 Db 711 AAGACCGTTGGGCAAAATTTCCACAATTAAGAGTGTGATGAGAGATGTTACACCAT 770
 QY 542 TGATGCTGTAAATTTGTTGAATCTTGCACATGAAGATTTGTTGCACACATGTCACCCC 601
 Db 771 TGATGCTGTAAATTTGTTGAATCTTGCACATGAAGATTTGTTGCACACATGTCACCCC 830
 QY 602 ACAACATCTTGTTTGAAACAGGAATTTCTCTTCCAAAGGGGCTTACMACCGCATATTA 661

Db 831 ACACATCTTGTGTAAGAGGAAATCTCTCTCAAGGGGGCTTTACACCGCATATTA 890
Qy 662 CTGCTTCAGTCTCTCAAAAAGAGATCCACAGGAGGACCTTGTCACGTGTAACAAG 721
Db 891 CTGCTTCAGTCTCTCAAAAAGAGATCCACAGGAGGAGCGCTGTGTACAGCTTAACAAG 950
Qy 722 TGGAGTAAAGATTTTCTTGGGACTGATAGTCTCTCATGATAGACGAAGAA 778
Db 951 TGGAGTAAAGATTTTCTTGGGACTGATAGTCTCTCATGATAGACGAAGAA 1007
RESULT 3 751 bp mRNA linear EST 07-MAR-2003
B0516584/c EST623999 Generation of a set of potato cDNA clones for microarray
LOCUS analyses mixed potato tissues Solanum tuberosum cDNA clone STM1Y71
DEFINITION 3' end, mRNA sequence.
ACCESSION B0516584
VERSION B0516584.1 GI:21375453
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanales; Solanaceae; Solanum.
Qy 1 (bases 1 to 751)
Db 1 Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
Karamecheva, S.A.
REFERENCE Generation of a set of potato cDNA clones for microarray analyses
AUTHORS Unpublished (2002)
OTHER ESTs: EST623998
TITLE The Institute for Genomic Research
JOURNAL 9712 Medical Center Dr, Rockville, MD 20850, USA
COMMENT Email: potato-array@ligr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T7
FEATURES
SOURCE Location/Qualifiers
1..751
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cuiivar="Kennebec or Binje"
/db_xref="taxon:4113"
/clone="STM1Y71"
/issue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/note="Vector: plasmid SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: combination of untreated and phytochrome
inhibitors-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating
eyes, tubers, or roots."

Db 631 GTTGAATCTTGCACTGAAGATTTGTGAGCAACTGTGACCCCAACATCTTTTGG 572
Qy 618 AACAGAAATCTCTCTTCCAAAGGGGCTTACAACGCAATATTAACCTTCAGCTC 677
Db 571 AACAGAAATCTCTCTTCCAAAGGGGCTTACAACGCAATATTAACCTTCAGCTC 512
Qy 678 AAAAGAGATCCACAGGAGGCACTTGTGACGTGTAAAGTGAAGTAAAGATT 737
Db 511 AAAAGAGATCCACAGGAGGCGCTGTGTACGCTTAACAAGTGAAGTAAAGATT 452
Qy 738 TTTCTTGGAGCTGATAGTCTCTCTCATGATAGCAAGAAAGAGTGTGAAAGAA 797
Db 451 TTTCTTGGAGCTGATAGTCTCTCTCATGATAGCAAGAAAGAGTGTGATGATG 392
Qy 798 GCTGATTTTAAATGACCTGAGCTGTGATGATGATGATGATGATGATGATGATG 857
Db 391 GCTGATTTTAAATGACCTGAGCTGTGATGATGATGATGATGATGATGATGATG 332
Qy 858 AATGCACTGACAAAGCTTGAAGCATTCAGTCTTCAATGACCAAGATTTTATGAGCTT 917
Db 331 AATGCACTGACAAAGCTTGAAGCATTTACTAGCTTCAATGACCAAGATTTTATGAGCTT 272
Qy 918 CCTAGAACAACTCAAAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 977
Db 271 CCTAGAACAACTCAAAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 212
Qy 978 TCTTATGACATCAGAAATATATATTTCCAGTGTGCTGTAATGCTGCACTGTTGCCG 1037
Db 211 TCTTATGACATCAGAAATATATATTTCCAGTGTGCTGTAATGCTGCACTGTTGCCG 152
Qy 1038 GCTCCTCTGAGAAATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1097
Db 151 GCTCCTCTGAGAAATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 92
Qy 1098 ATAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1157
Db 91 ATAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 32
Qy 1158 ATCTTTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1188
Db 31 ATCTTTTCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1
RESULT 4
LOCUS CK253198 915 bp mRNA linear EST 30-JUL-2004
DEFINITION EST736835 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POC351 5' end, mRNA sequence.
ACCESSION CK253198
VERSION CK253198.1 GI:39807898
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanales; Solanaceae; Solanum.
Qy 1 (bases 1 to 915)
Db 1 Buell, C.R., Hart, A., Zismann, V., Karamecheva, S.A. and Baker, B.
REFERENCE Generation of ESTs from potato callus tissue
AUTHORS Unpublished (2003)
JOURNAL Other ESTs: EST736836
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@ligr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
SOURCE Location/Qualifiers
1..915
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cuiivar="Kennebec"

/db_xref="taxon:4113"
/clone="POCC351"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI, Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match

Best Local Similarity 54.4%; Score 690.8; DB 7; Length 915;
Matches 695; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
OY 2 TGCAGAAATGAGAGCTCTCAATCAACACAGCTGATGATGGCATCTTCCTCGTATG 61
Db 214 TGCAGAAATGAGAGCTCTCAATCAACACAGCTGATGATGGCATCTTCCTCGTATG 61
OY 62 TGATGTTCTTAAGGAGCTGCTCTCAAGTCACATCACTTTGGAGGGCAATGATCAT 121
Db 274 TGATGTTCTTAAGGAGCTGCTCTCAAGTCACATCACTTTGGAGGGCAATGATCAT 121
OY 122 GCCAAATTTGAAGCTCTCTATCACTACCTGCTGCTGTAGACATACCGGAGCGCAT 181
Db 334 GCCAAATTTGAAGCTCTCTATCACTACCTGCTGCTGTAGACATACCGGAGCGCAT 181
OY 182 ATTGAATCTTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 241
Db 394 ATTGAATCTTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 241
OY 242 TACAAACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301
Db 454 TACAAACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 301
OY 302 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
Db 514 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
OY 362 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
Db 574 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
OY 422 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
Db 634 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
OY 482 AAGACCGTTGAGAGAAATTTCCAAATTTGAAGAGTGTGATGAGACATGTTACCACT 541
Db 694 AAGACCGTTGAGAGAAATTTCCAAATTTGAAGAGTGTGATGAGACATGTTACCACT 541
OY 542 TGATGCTGTTAAGTTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 601
Db 754 TGATGCTGTTAAGTTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 601
OY 602 ACAACATCTGTTTGAAGAGAAATTTCTCTTCCAAAGGGGCTTACCAACCCATATTA 661
Db 814 ACAACATCTGTTTGAAGAGAAATTTCTCTTCCAAAGGGGCTTACCAACCCATATTA 661
OY 662 CTGCTTCCAGTCTCTCAAAAGAGATCCACAGGAGGAGCAT 703
Db 874 CTGCTTCCAGTCTCTCAAAAGAGATCCACAGGAGGAGCAT 703
```

RESULT 5

CK242513/c

LOCUS

DEFINITION CK242513 790 bp mRNA linear EST 30-JUL-2004
EST726150 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POCA005 3' end, mRNA sequence.

ACCESSION

CK242513

VERSION

CK242513.1

KEYWORDS

EST.

SOURCE

Solanum tuberosum (potato)

ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 790)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from potato callus tissue
Unpublished (2003)
Other ESTs: EST726147 EST726148 EST726149
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: GTA ATA CGA CTC ACT ATA GGG C.
Location/Qualifiers
1..790

FEATURES

source

/organism="Solanum tuberosum"
/mol_type="mRNA"
/culivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCA005"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI, Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 54.3%; Score 689.6; DB 7; Length 790;
Best Local Similarity 95.9%; Pred. No. 1.8e-179;
Matches 727; Conservative 0; Mismatches 9; Indels 22; Gaps 1;

```
OY 497 GAATTTCCAAATTTGAAGAGTCTGATGAGCATGTTACCACTTATGCTGTTAAGTT 556
Db 790 GAATTTCCAAATTTGAAGAGTCTGATGAGCATGTTACCACTTATGCTGTTAAGTT 556
OY 557 TGTGAATCTTTCAGTGAAGATTTGTGACAGCATGTCACCCCAACATCTGTTT 616
Db 730 TGTGAATCTTTCAGTGAAGATTTGTGACAGCATGTCACCCCAACATCTGTTT 616
OY 617 GAACAGAAATCTCTTCCAAAGGGGCTTACACCGCATTAATTACTGCTTCCAGTCT 676
Db 670 GAACAGAAATCTCTTCCAAAGGGGCTTACACCGCATTAATTACTGCTTCCAGTCT 676
OY 677 CAAGAAGATCCACAGAGAGGCACTGTGTACAGCTGTAACAAGTAAAGATT 736
Db 610 CAAGAAGATCCACAGAGAGGCGCTGTGTGTACAGCTGTAACAAGTAAAGATT 736
OY 737 TTTTCTGGAGCTGATGATGCTCTCATGATGACGAAGAAAGAGTCTTGTGATG 796
Db 550 TTTTCTGGAGCTGATGATGCTCTCATGATGACGAAGAAAGAGTCTTGTGATG 796
OY 550 TTTTCTGGAGCTGATGATGCTCTCATGATGACGAAGAAAGAGTCTTGTGATG 796
Db 490 TGTGATTAATTAATGACACCGTATGCTGTGATGATGACGAAGAGTGTGAAAAGA 431
OY 490 TGTGATTAATTAATGACACCGTATGCTGTGATGATGACGAAGAGTGTGAAAAGA 431
OY 857 AATGACATCCGACAGCTTGAAGCATTAAGCTTCAATGACAGATTTTATGAGCT 916
Db 430 AATGACATCCGACAGCTTGAAGCATTAAGCTTCAATGACAGATTTTATGAGCT 916
OY 917 TCTTAGAACAACCTGAAGATTAGTTAGTAAGAGCCATGGAAGTACCCGAATCTT 976
Db 370 TCTTAGAACAACCTGAAGATTAGTTAGTAAGAGCCATGGAAGTACCCGAATCTT 976
OY 977 TTCTTAGCATCAGAGATTAATTAATTCATGTTGTGCTGTAAGTCTGCACTGTTGCC 1036
Db 310 TTCTTAGCATCAGAGATTAATTAATTCATGTTGTGCTGTAAGTCTGCACTGTTGCC 251
OY 1037 GGCTCTCTCTGAGAA-----TCATTGTCAATTTCTTGAATCTG 1074
```

Db 250 GGGTCTCTCTGAGAAAGGAAAGAGATTCCTCTTCATTTGATCTTGACCTG 191
QY 1075 AATATGTGATTAACCAAGATATAGACTGTAGTATCATCTTTCTTCATGTTGA 1134
Db 190 AATATGTGATTAACCAAGATATAGACTGTAGTATCATCTTTCTTCATGTTGA 131
QY 1135 TTAGATATTAATCAAGATATATCTTTCACTTAATAATATGGAACAATAAGCTT 1194
Db 130 TTGATATATTATCCGATGATATATCTTTCACTTAATAATATGGAACAATAAGCTT 71
QY 1195 TGCACGCTACCAAGGCTCTCTGATTCCTGAAGTCT 1232
Db 70 TGCACGCTACCAAGGCTCTCTGATTCGAAGATTT 33

RESULT 6
LOCUS CK250612 920 bp mRNA linear EST 30-JUL-2004
DEFINITION EST734249 potato callus cDNA library, normalized and full-length
Solanium tuberosum cDNA clone POCBH29 5' end, mRNA sequence.
ACCESSION CK250612 GI:39802848
VERSION CK250612.1
KEYWORDS EST.
SOURCE Solanium tuberosum (potato)
ORGANISM Solanium tuberosum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanium.
1 (bases 1 to 920)
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
source
1..920
/organism="Solanium tuberosum"
/mol_type="mRNA"
/cultiyar="Kennebec"
/db_xref="taxon:4113"
/clone="POCBH29"
/issue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanium tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN
Query Match 52.1% Score 662.6; DB 7; Length 920;
Best Local Similarity 99.3%; Pred. No. 5.6e-172;
Matches 676; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 2 TGCAAAATGAGGCTCTCATACACAACCTGATGATGGCATCTTCCTCCGTATGG 61
Db 241 TGCAAAATGAGGCTCTCATACACAACCTGATGATGGCATCTTCCTCCGTATGG 300
QY 62 TGATGTTCTTAAGGAGTGTCTCTCAAGTCACATACATTTGGAGGCAATGTGAT 121
Db 301 TGATGTTCTTAAGGAGTGTCTCTCAAGTCACATACATTTGGAGGCAATGTGAT 360
QY 122 GCCAATTTGAAGCTCTCATACACACTGCTGCTGTAGCATACCGGAGGCGAT 181
Db 361 GCCAATTTGAAGCTCTCATACACACTGCTGCTGTAGCATACCGGAGGCGAT 420
QY 162 ATTGAATCTTTACTCTGTTGATGATTTCAACCTCTTATGACACTTTATTTGACGA 241

Db 421 ATTGAATCTTTACTCTGTTGATGATTTCAACCTCTTATGACACTTTATTTGACGA 480
QY 242 TCAACCAAGCTCTTATGGAATATCAACTACGAAGAGACCGAGTGTATTTGGGGTGA 301
Db 481 TCAACCAAGCTCTTATGGAATATCAACTACGAAGAGACCGAGTGTATTTGGGGTGA 540
QY 302 GTTGTACCTGCTGGGCGACGCAAAATCTCAAGATGAGTACATCTTTCCGGGA 361
Db 541 GTTGTACCTGCTGGGCGACGCAAAATCTCAAGATGAGTACATCTTTCCGGGA 600
QY 362 GTTGTACCAAGCTCTTCAAGAAATGTTGAGCATTAATATGCTCTGCTGTTATGAGA 421
Db 601 GTTGTACCAAGCTCTTCAAGAAATGTTGAGCATTAATATGCTCTGCTGTTATGAGA 660
QY 422 GGTACTTAATCTGAGCTTGACATGTTTGTATGAGAAAGATTTATTTGAAACGGTCT 481
Db 661 GGTACTTAATCTGAGCTTGACATGTTTGTATGAGAAAGATTTATTTGAAACGGTCT 720
QY 482 AAGACCGTTGGTCAGAAATTTCCAAATTTGAAGTGTGATGAGCATGTTACCAAT 541
Db 721 AAGACCGTTGGTCAGAAATTTCCAAATTTGAAGTGTGATGAGCATGTTACCAAT 780
QY 542 TGATGCTGTTAAGTTGTTGAATCTTGCACTGGAAGATTTGTCAGCAACTGTACCCC 601
Db 781 TGATGCTGTTAAGTTGTTGAATCTTGCACTGGAAGATTTGTCAGCAACTGTACCCC 840
QY 602 ACAACATCTTGTGTTGAACGGAATTCCTCTTCCAAAGGGGCTTTACACCGCATATTA 661
Db 841 ACAACATCTTGTGTTGAACGGAATTCCTCTTCCAAAGGGGCTTTAC-ACCGCATATTA 899
QY 662 CTGCTTCCAGTCTCTCAAAAG 682
Db 900 CTGCTTCCAGTCTCTCAAAAG 920

RESULT 7
CK249779/c 825 bp mRNA linear EST 30-JUL-2004
LOCUS EST734346 potato callus cDNA library, normalized and full-length
DEFINITION EST734346 potato callus cDNA clone POCBA88 3' end, mRNA sequence.
ACCESSION CK249779
VERSION CK249779.1 GI:39801235
KEYWORDS EST.
SOURCE Solanium tuberosum (potato)
ORGANISM Solanium tuberosum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanium.
1 (bases 1 to 825)
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST733415
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: GTA ATA CGA CTC ACT ATA GGG C.

FEATURES
source
1..825
/organism="Solanium tuberosum"
/mol_type="mRNA"
/cultiyar="Kennebec"
/db_xref="taxon:4113"
/clone="POCBA88"
/issue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

ORIGIN

supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."

Query Match 51.8%; Score 658.2; DB 7; Length 825;
Best Local Similarity 94.6%; Pred. No. 9e-171;
Matches 701; Conservative 0; Mismatches 18; Indels 22; Gaps 1;

```

QY 414 CATGAGAGAGTTACTTAATCTCGAGTTGACATGTTTATAGAGAAAAGTATTCATTGAA 473
DB 825 CATGAGAGAGTTACTTAATCTCGAGTTGACATGTTTATAGAGAAAAGTATTCATTGAA 766
QY 474 ACGGTTCTAAGACCGTTGTGACAGAAATTTCCAAATTGAAGTGTGATGAGACATGTT 533
DB 765 ACGGTTCTAAGACCGTTGTGACAGAAATTTCCAAATTGAAGTGTGATGAGACATGTT 706
QY 534 ACCACCATTTGATGCTGTAGTTGTTGAATCTTTCATCGAAGAGATTTGTTCAGCAACT 593
DB 705 ACCACCATTTGATGCTGTAGTTGTTGAATCTTTCATCGAAGAGATTTGTTCAGCAACT 646
QY 594 GTCAACCCCAACAATCTTGTGTTGAAACAGGAATTTCTCTCCAAAGGGGCTTACACCG 653
DB 645 GTCAACCCCAACAATCTTGTGTTGAAACAGGAATTTCTCTCCAAAGGGGCTTACACCG 586
QY 654 CATATTTACTGCTCCCTCAAGCTCAAAAGAGATCCACAGGAGGCACTTGTGTACAGT 713
DB 585 CATATTTACTGCTCCCTCAAGCTCAAAAGAGATCCACAGGAGGCACTTGTGTACAGT 526
QY 714 GTTACAAGTGAAGTAAAGATTTTCTTGGGACTGATAGGCTCCATGATAGACGA 773
DB 525 GTTACAAGTGAAGTAAAGATTTTCTTGGGACTGATAGGCTCCATGATAGACGA 466
QY 774 AGAAAGAGTGTCTTGTGATGCTGTGATTTTCAATGACACTGTAGCTTGTCACTA 833
DB 465 AGAAAGAGTGTCTTGTGATGCTGTGATTTTCAATGACACTGTAGCTTGTCACTA 406
QY 834 TATGCAAGGTGTTGAAAAGGAAATGCACTCGAAGCTTGAAGATTCATGACTTC 893
DB 405 TATGCAAGGTGTTGAAAAGGAAATGCACTCGAAGCTTGAAGATTCATGACTTC 346
QY 894 AATGACACGATTTTATGGGCTTCTTGAACCACTGAAAGATTTAGTGAAGAAGC 953
DB 345 AATGACACGATTTTATGGGCTTCTTGAACCACTGAAAGATTTAGTGAAGAAGC 286
QY 954 CCATGGAAGTAAACCGAATCTTCTTATGATCAGAGATATTTTCCCATGTTGCT 1013
DB 285 CCATGGAAGTAAACCGAATCTTCTTATGATCAGAGATATTTTCCCATGTTGCT 226
QY 1014 GGTGAAATGCTGCACTGTTGCCGGCTCTCTTGAGAA----- 1052
DB 225 GGTGAAATGCTGCACTGTTGCCGGCTCTCTTGAGAA----- 166
QY 1053 -TCATTTGTCAATCTTGTACTGTAAATTTGATGATCAACCAAGATTTAGCTGAGGTG 1111
DB 165 CTCATTTGTCAATCTTGTACTGTAAATTTGATGATCAACCAAGATTTAGCTGAGGTG 106
QY 1112 TATCATCTTTCTTTTCAATGTT 1132
DB 105 TATCATCTTTCTTTTCAATGTT 85

```

RESULT 8
CK242510/c 723 bp mRNA linear EST 30-JUN-2004
LOCUS EST726147 potato callus cDNA library, normalized and full-length
DEFINITION Solanum tuberosum cDNA clone PCCA005 3' end, mRNA sequence.
ACCESSION CK242510
VERSION CK242510.1 GI:39786205
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
1 (bases 1 to 723)
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karayancheva, S.A. and Baker, B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST726148 EST726149 EST726150
Contact: Robin Buell

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igrr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/.
Seq primer: TGT AAA ACG ACG GCC AGT.

FEATURES

Location/Qualifiers

```

1..723
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="PCCA005"
/tissue_type="callus"
/lab_host="DH10B-Roma"
/clone_lib="potato callus cDNA library, normalized and
full-length"
/note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

```

ORIGIN

Query Match 48.5%; Score 616.2; DB 7; Length 723;
Best Local Similarity 94.9%; Pred. No. 3.7e-159;
Matches 656; Conservative 0; Mismatches 13; Indels 22; Gaps 1;

```

QY 564 TCTTGACATGAGATTTGTTGACAGCACTGTACCCCAACAATCTGTTTGAACAG 623
DB 723 TCTTGACATGAGATTTGTTGACAGCACTGTACCCCAACAATCTGTTTGAACAG 664
QY 624 AATTCTCTCTCCAAAGGGGGCTTACAAACCGATATTAATGCTTCCAGTCTCCAAAGA 683
DB 663 AATTCTCTCTCCAAAGGGGGCTTACAAACCGATATTAATGCTTCCAGTCTCCAAAGA 604
QY 684 GAGATCCAGAGGAGGCACTTGTGAGCTGTAAAGAGTAAAGTAAAGATTTTCTT 743
DB 603 GAGATCCAGAGGAGGCGCTTGTGAGCTGTAAAGAGTAAAGTAAAGATTTTCTT 544
QY 744 GGAATCTGATGCTCTCTCATGATAGACGAGAAAGAGTCTTGTGATGCTGCT 803
DB 543 GGAATCTGATGCTCTCTCATGATAGACGAGAAAGAGTCTTGTGATGCTGCT 484
QY 804 ATTACAAATGCACTGTGACCTTGTCAAGTATATGCGAAGGTGTTGAAAAGAAATGCA 863
DB 483 ATTACAAATGCACTGTGACCTTGTCAAGTATATGCGAAGGTGTTGAAAAGAAATGCA 424
QY 864 CTGCAACACTGGAAGACATTCATGACTTCAATGAGACCAAGATTTTATGGGCTTCTAGG 923
DB 423 CTGCAACACTGGAAGACATTCATGACTTCAATGAGACCAAGATTTTATGGGCTTCTAGG 364
QY 924 AACCACTCAAGATTAAGTTAGTGAAGCGCAATGAGAGATCCCGAATCTTTTCTTAT 983
DB 363 AACCACTCAAGATTAAGTTAGTGAAGCGCAATGAGAGATCCCGAATCTTTTCTTAT 304
QY 984 GCATCAGAGATATTAATTCAGATGTTGCGTGAAGATGCTGACATGTTGCCGGCTCCT 1043
DB 303 GCATCAGAGATATTAATTCAGATGTTGCGTGAAGATGCTGACATGTTGCCGGCTCCT 244
QY 1044 CTCTGAGAA-----TCATTTGTATTTCTTGTATCTGTAATATTTG 1081
DB 243 TTTTGAGAAAGGAGAAAGAGATTTCTCTTCTCATATTTGTCAATCTTGTACTGTAAATTTG 184
QY 1082 TGATTTCAACCAAGATATTAAGTGTAGTGTATCATCTTTCTTATGATGATTAAGATA 1141
DB 183 TGATTTCAACCAAGATATTAAGTGTAGTGTATCATCTTTCTTATGATGATTAAGATA 124

```


QY 1142 TTATCAGATGATATATCTTTCAGCTAATAATTAAGAAACAATAGCTTTGACGC 1201
| | | | |
DB 123 TTATCCGATGATAAATCTTTTCAGCTAATAATTAAGAAACAATAGCTTTGACGC 64
| | | | |
QY 1202 TCACCAAGTCTCTGTAATTTCTGAAGTTCT 1232
| | | | |
DB 63 TCACCAAGTCTCTGTAATTTCTGAAGTTCT 33
| | | | |
RESULT 9 708 bp mRNA linear EST 30-JULY-2004
CK251510/c 708 bp mRNA linear EST 30-JULY-2004
LOCUS EST35147 potato callus cDNA library, normalized and full-length
DEFINITION Solanum tuberosum cDNA clone POCB96 3' end, mRNA sequence.
ACCESSION CK251510
VERSION CK251510.1 GI:39804606
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
REFERENCE Solanum tuberosum; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 708)
Buell, C.R., Hart, A., Zismann, V., Karanymcheva, S.A. and Baker, B.
AUTHORS Generation of ESTs from potato callus tissue
TITLE Unpublished (2003)
JOURNAL Contact: Robin Buell
COMMENT The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via <http://genome.arizona.edu/orders/>.
Seq primer: GTA ATA CGA CTC ACT ATA GGC C.

FEATURES

source
1. 708
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCB96"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_id="potato callus cDNA library, normalized and
full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 47.8%; Score 607.6; DB 7; Length 708;
Best Local Similarity 95.4%; Pred. No. 8.9e-157;
Matches 645; Conservative 0; Mismatches 9; Indels 22; Gaps 1;
QY 579 TTGTGTGAGCACTGTGACCCACACATCTTTTGAACAGGAATCTCTCTCCAA 638
| | | | |
DB 708 TTGTGTGAGCGAGCTGACCCACACATCTTTTGAACAGGAATCTCTCTCCAA 649
| | | | |
QY 639 GGGGGCTTACAAACGCAATTAATCTGCTGAGCTCTCAAAAGAGATCCACAGGAG 698
| | | | |
DB 648 GGGGGCTTACAAACGCAATTAATCTGCTGAGCTCTCAAAAGAGATCCACAGGAG 589
| | | | |
QY 699 GCACTTGTGAGCTGTGAACAAGTGAAGTAAGATTTTCTTGGAGCTGATAGTCT 758
| | | | |
DB 588 GCGCTGTGTCAGCTGTGAACAAGTGAAGTAAGATTTTCTTGGAGCTGATAGTCT 529
| | | | |
QY 759 CCTCATGATTAAGCAAGAAAGAGTCTTCTGTGATGTGCTGTAATTTCAATGCAACT 818
| | | | |
DB 528 CCTCATGATTAAGCAAGAAAGAGTCTTCTGTGATGTGCTGTAATTTCAATGCAACT 469
| | | | |
QY 819 GTAGCTTGTGATGATGCAAGAGTGTGTTGAAGAAAGAAATGCACTGCAAGAGCTTGA 878
| | | | |
DB 468 GTAGCTTGTGATGATGCAAGAGTGTGTTGAAGAAAGAAATGCACTGCAAGAGCTTGA 409
| | | | |

QY 879 GCATTCAGTCTGATGATGACCAAGATTTTATAGGCTTCTTAAGAAACAATCAAGATT 938
| | | | |
DB 408 GCATTTACTAGCTTCAATGAGCAAGACTTTATAGGCTTCTTAAGAAACAATCAAGATT 349
| | | | |
QY 939 AAGTTAGTAAGACCGCATGGAAGGTACCCGAATCTTTTCTTATGATCAGAGATTT 998
| | | | |
DB 348 AAGTTAGTAAGACCGCATGGAAGGTACCCGAATCTTTTCTTATGATCAGAGATTT 289
| | | | |
QY 999 ATTCCCATTTTCTGCTGTAATAGCTGACCTGGTTCGGGCTCTCTGAGAA----- 1052
| | | | |
DB 288 ATTCCCATTTTCTGCTGTAATAGCTGACCTGGTTCGGGCTCTCTGAGAAAGGAA 229
| | | | |
QY 1053 -----TCATTTGTCACTTGTGATGTAATATTTGATTTCAACCAAGA 1096
| | | | |
DB 228 AAGAGATTCTCTTCTCATTTGTCTCTTGTGATGTAATATTTGATTTCAACCAAGA 169
| | | | |
QY 1097 TATAGACTGAGTGTATGATCATCTTTTCTTATGATTTATGATTTATCATGATTA 1156
| | | | |
DB 168 TATAGACTGAGTGTATGATCATCTTTTCTTATGATTTATGATTTATCATGATTA 109
| | | | |
QY 1157 TATCTTTTCACTAATTAATTAAGAAACAATAGCTTTGACAGCTCACCAAGTGCTCC 1216
| | | | |
DB 108 TATCTTTTCACTAATTAATTAAGAAACAATAGCTTTGACAGCTCACCAAGTGCTCC 49
| | | | |
QY 1217 TGTATCTGAAAGTTCT 1232
| | | | |
DB 48 TGTATCTGAAAGTTCT 33
| | | | |

RESULT 10

BI927946 732 bp mRNA linear EST 18-OCT-2001
LOCUS EST347823 tomato flower, 3 - 8 mm buds Lycopersicon esculentum cDNA
DEFINITION clone CT0B24AI 5' end, mRNA sequence.
ACCESSION BI927946
VERSION BI927946.1 GI:16238481
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 732)
van der Hoeven, R.S., Bezzerides, J.L., Karanymcheva, S.A., Tsai, J.,
AUTHORS Utechtack, T., Van Aken, S., Roming, C.M., Nierman, W., Fraser, C.M.,
Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
TITLE Generation of ESTs from tomato flower tissue, 3-8 mm buds (2001)
JOURNAL Contact: CUGI
COMMENT Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.

FEATURES

source
1. 732
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CT0B24AI"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/clone_id="tomato flower, 3 - 8 mm buds"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research. Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN

Query Match 47.2%; Score 599.4; DB 4; Length 732;
 Best Local Similarity 97.4%; Pred. No. 1.7e-154;
 Matches 609; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

QY 2 TGC AAAAATGAGAGCTCTCAATCAACACACCTGATGATGGACATCTTCATCCGATG 61
DB 108 TGC AAAAATGAGAGCTCTCAATCAACACACCTGATGATGGACATCTTCATCCGATG 167
QY 62 TGAATGTTCTTAAGGAGGAGTGTCTCTCAAGTGCACATCTTGGAGGGCAATATG 121
DB 168 TGAATGTTCTTAAGGAGGAGTGTCTCTCAAGTGCACATCTTGGAGGGCAATATG 227
QY 122 GCCAAATTTGAAGCCTCTCTATCACTACCACTGCTGCTGCTGATGACATCCGGAGG 181
DB 228 GCCAAATTTGAAGCCTCTCTATCACTACCACTGCTGCTGCTGATGACATCCGGAGG 287
QY 182 ATGGAATCTTTAACCCTGATGATGATGATTTCAACCTCTTATGACACTTTATTTGAC 241
DB 288 ATGGAATCTTTAACCCTGATGATGATGATTTCAACCTCTTATGACACTTTATTTGAC 347
QY 242 TACAAACAGTCTTATGGAATCAAACTAGCAAGAGAGCCAGGCTGATTTGGGGTAA 301
DB 348 TACAAACAGTCTTATGGAATCAAACTAGCAAGAGAGCCAGGCTGATTTGGGGTAA 407
QY 302 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
DB 408 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 467
QY 362 GTTGTACCAAGTCTTACAAAGAAATGTTGAGCAATTAATGCTCTGCTGCTGCTGCTG 421
DB 468 GTTGTACCAAGTCTTACAAAGAAATGTTGAGCAATTAATGCTCTGCTGCTGCTGCTG 527
QY 422 GGTACTAATCTGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
DB 528 GGTACTAATCTGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 587
QY 482 AAGACCGTGTGTCAGAAATTTCCAAATTTGAAGTGTGATGAGATGATGATGATGAT 541
DB 588 AAGACCGTGTGTCAGAAATTTCCAAATTTGAAGTGTGATGAGATGATGATGATGAT 647
QY 542 TGATGCTTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
DB 648 TGATGCTTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 707
QY 602 ACAACATCTTGTGTTGAACAGGAT 626
DB 708 ACAACATCTTGTGTTGAACAGGAT 732

```

RESULT 11
 CK242511 957 bp mRNA linear EST 30-JUL-2004
 LOCUS EST726148 potato callus cDNA library, normalized and full-length
 DEFINITION Solanum tuberosum cDNA clone POC4005 5' end, mRNA sequence.
 ACCESSION CK242511
 VERSION CK242511.1 GI:39786207
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Bukarjoca; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 957)
 Buell, C.R., Hart, A., Zismann, V., Karameycheva, S.A. and Baker, B.
 Generation of ESTs from potato callus tissue
 Unpublished (2003)
 Other ESTs: EST726147 EST726149 EST726150
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@igr.org
 Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>
 Seq primer: ATT TAG GTG ACA CTA TAG.
 Location/Qualifiers
 1. 957

FEATURES

SOURCE

/organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POC4005"
 /issue_type="callus"
 /lab_host="DH10B-Tona"
 /clone_lib="potato callus cDNA library, normalized and
 full-length"
 /note="Vector: PCWVspore6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Solanum tuberosum var.
 Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 46.8%; Score 595; DB 7; Length 957;
 Best Local Similarity 99.0%; Pred. No. 3e-153;
 Matches 609; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

QY 2 TGC AAAAATGAGAGCTCTCAATCAACACACCTGATGATGGACATCTTCATCCGATG 61
DB 341 TGC AAAAATGAGAGCTCTCAATCAACACACCTGATGATGGACATCTTCATCCGATG 400
QY 62 TGAATGTTCTTAAGGAGGAGTGTCTCTCAAGTGCACATCTTGGAGGGCAATATG 121
DB 401 TGAATGTTCTTAAGGAGGAGTGTCTCTCAAGTGCACATCTTGGAGGGCAATATG 460
QY 122 GCCAAATTTGAAGCCTCTCTATCACTACCACTGCTGCTGCTGATGACATCCGGAGG 181
DB 461 GCCAAATTTGAAGCCTCTCTATCACTACCACTGCTGCTGCTGATGACATCCGGAGG 520
QY 182 ATGGAATCTTTAACCCTGATGATGATGATGATTTCAACCTCTTATGACACTTTATTTGAC 241
DB 521 ATGGAATCTTTAACCCTGATGATGATGATTTCAACCTCTTATGACACTTTATTTGAC 580
QY 242 TACAAACAGTCTTATGGAATCAAACTAGCAAGAGAGCCAGGCTGATTTGGGGTAA 301
DB 581 TACAAACAGTCTTATGGAATCAAACTAGCAAGAGAGCCAGGCTGATTTGGGGTAA 640
QY 302 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
DB 641 GTTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 700
QY 362 GTTGTACCAAGTCTTACAAAGAAATGTTGAGCAATTAATGCTCTGCTGCTGCTGATG 421
DB 701 GTTGTACCAAGTCTTACAAAGAAATGTTGAGCAATTAATGCTCTGCTGCTGCTGATG 760
QY 422 GGTACTAATCTGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
DB 761 GGTACTAATCTGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 820
QY 482 AAGACCGTGTGTCAGAAATTTCCAAATTTGAAGTGTGATGAGATGATGATGATGAT 541
DB 821 AAGACCGTGTGTCAGAAATTTCCAAATTTGAAGTGTGATGAGATGATGATGATGAT 880
QY 542 TGATGCTTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
DB 881 TGATGCTTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 939
QY 602 ACAACATCTTGTGTT 616
DB 940 CACACATCTTGGTTT 954

```

RESULT 12

CNS0A31G 1301 bp mRNA linear HTC 06-FEB-2004
 LOCUS CNS0A31G
 DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 GSLPBG902805 of Hormone Treated Callus of strain col-0 of
 Arabidopsis thaliana (thale cress) ..

ACCESSION BX828455
VERSION BX828455.1 GI:42460819
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 1301)
REFERENCE Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
TITLE Unpublished
JOURNAL 2 (bases 1 to 1301)
REFERENCE Genoscope.
AUTHORS Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 131 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES
Source Location/Qualifiers
1..1301
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="COL-0"
/db_xref="taxon:3702"
/clone="GSLTGH90ZB05"
/issue_type="Hormone Treated Callus"
/plasmid="PCWVSPORT_6"
1..1301
/gene="At4g22930"
ORIGIN
Query Match 46.1%; Score 585.4; DB 3; Length 1301;
Best Local Similarity 72.9%; Pred. NO. 1.5e-150;
Matches 754; Conservative 0; Mismatches 281; Indels 0; Gaps 0;
3 GCAAAATGAGGCTCTCAATACACACAACTGATGATGGGACTCTTCATCCCGATGAT 62
187 GTAAGAAAGGCTCAACATCACTACCTATGATGATGGGACTCTTCATCCCGATGAT 246
63 GATGTTCTTAAGGAGTGTCTCTACAGTGCACATCACTTTGGAGGGCAATAGTCATG 122
247 GATCTTCTCATGCTGTGTTGTTCCCAACAGTGCAGATATTTAAGAGGCAATGTGATG 306
123 CCAAAATTTGAAGCTCTCTATCACTACCACTGCTGCTGTGATGATACCGGAGCGCAT 182
307 CCGAATCGAAGCTCTCTGTCACCTCCACTGCACTGCACTTATTTACCGAAATTTATC 366
183 TTGAATCTTAACTGTTAGTAGATTTCAACCTCTTATGACACTTATTTAGACAGAT 242
367 ATGAAAGCTTTCATCTGAAGATGACTTTGATTCATTAAGACCTTTATTTAGTAC 426
243 ACAACCAAGTCTTATGAAATCAAACTACAGAGAGAGACCGATGATTTGGGGTGAAG 302
427 AAAAATCTACCTGAAGATCAGGTTTCCAAAGGAAAGTGGTGTGTTATGCGGTGAAG 486
303 TTGATCCTGCTGTCACACGAAATTTCTCAAGATGAGTGAATGATCTTTTGGGAAG 362

DB 487 CTGTACCTCCCGAGCAACACCACTCTCAAGATGATGTCACGATCTCTTTGGAAA 546
363 TGTTTACAGATTTCTACAGAAATGTTGAGCATATATGCTCTGCTGTTATGAGAG 422
547 TGTCTACCGGATCTAGAAAGATGTCGCAACAAATGCTCTTGTGTTATGAGGAG 606
423 GTTACTAATCTGAGGTTGATCATGTTGATGAGAAAGGATTTCAATTGAAACGTTCT 482
607 GTACAGATCCGATATCAATGATCTTTGACCGGAGAAATCTTCAATTGAGACGTTCT 666
483 AGACCTTGATGAGAAATTTCCAAATTTGAAGTGTGATGAGAGATGTTACCACTT 542
667 CAGCTCTAATCCAAAGCGCTCCGACGTGAAGTGTGATGAGACATCAACACATG 726
543 GATGCTGTTAAGTTGTTGATTTGATCTTCACTGAAAGATTTGTCACCACTGCA 602
727 GATGCTGTAATTTGTTGATCTTCAAGAGAGGATCTGTGATGATCAAAATGAC 786
603 CACATCTGTTTGTGAAGAGAAATCTCTCTTCCAAAGGGGCTTACAAACCGCATTA 662
787 CAACATCTCTTCTCAACAGAAAGCGCTTTTCCAAAGTGTGATTAACCTCAACTAC 846
663 TGCCTTCAAGTCTCTCAAAAGAGATCCACAGGAGGCACTTGTCACTGTAACAGT 722
847 TGTCTTCCCGTCTTCAAAAGAAATATACATGAGAGGCACTTGTCAAGCTGTA 906
723 GGAAGTAAAGATTTTCTTGGAGCTGATGATGCTCTCATATGATAGCAAGAAAG 782
907 GGAAGCAAGAAATTTCTTCCGACAGATAGGCTCCACATTAACGAGATAGAAAG 966
783 TGTCTTGGAGATGCTGTGATTTTCAATGACACCTGTCATATATGAGAG 842
967 TCATCTCGGAGATGCTGTGATTTTACAGGCTTCCATCTTGTCTTATGAGCA 1026
843 GTGTTGAAAGAAATGACATCGACACAGCTGAAACATTAAGCTTCAATGAGCA 902
1027 GTCTTGTATAGGCGGCTGCGCTGACATATGGAAGCTTCAACAGTTCAATGAG 1086
903 GATTTTATGAGGCTCTCTAGAACTCAAGATTAAGTTAGTAAGAGCCATGAG 962
1087 GATTTTATGAGGCTCTCTAGAACTCGTCAAGATCACTGAAAGATCTCTTGA 1146
963 GTACCGAATCTTTTCTTATGATCAAGATATTAATTCAGTTGCTGAGAAAG 1022
1147 GTTCCAGACGTTTCAACTTCCATTCGAGAGATGCTCTTATGTTGCTGAGAA 1206
1023 CTCGACTGTTGCG 1037
1207 CTTCATGCAACCG 1221
RESULT 13
CNS03385 1294 bp mRNA linear HTC 06-FEB-2004
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTGH93ZB09 of flowers and buds of strain col-0 of Arabidopsis
thaliana (thale cress).
ACCESSION BX827082
VERSION BX827082.1 GI:42461050
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
1 (bases 1 to 1294)
REFERENCE Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1294)
AUTHORS Genoscope.
TITLE Direct Submision
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Crnaud C., Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Piprap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full Length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
source Location/Qualifiers
1..1294
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLT932B09"
/tissue_type="Flowers and buds"
/plasmid="pCMVSPORT_6"
1..1294
/gene="At4g22930"

Query Match 45.9%; Score 583.8; DB 3; Length 1294;
Best Local Similarity 72.8%; Pred. No. 4.2e-150;
Matches 753; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

QY 3 GCAAAATGAGCTCTCAATCAACAACTGATGATTTGATCTCCGTGATGCT 62
DB 200 GTAGAGTGAAGCTCAATCACTCAACCTGATGATTTGATCTCCGTGACGGT 259
QY 63 GATGTTCTTAAAGGAGTGTCTCTCAAGTGCACATCACTTGGAGGGCAATGCTG 122
DB 260 GATCTCTTCAATGCTGTGTTCCCAACAGTGCAGATTTTGAAGACGATTTGAGT 319
QY 123 CCAATTTGAAGCTCTCAATCACTCACTGCTGCTGTGATACATACCGGAGCGATA 182
DB 320 CCGAATCTGAAGCTCTCTGACCTCCATGCACTGATTTATTAACCGAAATTTATC 379
QY 183 TTGAATCTTAACTGTTGATGATGATTTCAACCTCTTATGACCTTTATTTGACAGT 242
DB 380 ATGAAGCTTTGCACTGATGATGATTTGATTTGATTTGATTTGATTTGATTTG 439
QY 243 ACAACCAAGTCTTGAATCAAACTAGCAAGAGAGAGCAGATGATTTGGGTTGAAG 302
DB 440 AAAAAGCTTAACTGAAGATCAAGGCTTGAAGGAAAGTGATGTTTATGCGGTGAAG 499
QY 303 TTGTACCTGCTGCTGCTCAAGCAATTTCTCAAGATGAGTCACTGATTTTGGGAG 362
DB 500 CTGTACCTGCTGCTGCTCAAGCACTCTCAAGATGAGTCACTGATTTTGGGAG 559
QY 363 TGTTCACAGTTCTTACAAAGATGTTGATGATGATGATGATGATGATGATGAT 422
DB 560 TGTTCACAGTTCTTACAAAGATGTTGATGATGATGATGATGATGATGATGAT 619
QY 423 GTTACTAATCTGAGGTTGATGATGATGATGATGATGATGATGATGATGATGAT 482
DB 620 GTTACTAATCTGAGGTTGATGATGATGATGATGATGATGATGATGATGATGAT 679
QY 483 AAGACGTTGCTGAGGTTGATGATGATGATGATGATGATGATGATGATGATGAT 542
DB 680 CAGCTCTAATCTCAAGGCTTCCGACGCTGAAGATGATGATGATGATGATGATGAT 739

QY 543 GATGCTTTAAGTTTGTGATTAATCTTCACTGAAGATTTGTCAGCAACTGCCCA 602
DB 740 GATGCTTTAAGTTTGTGATTAATCTTCACTGAAGATTTGTCAGCAACTGCCCA 799
QY 603 CAACATCTTGTGTTGAACAGAAATCTCTTCAAGAGGAGGCTTCAACCGCATTAAT 662
DB 800 CAACATCTTGTGTTGAACAGAAACGCTTTTCAAGAGGATTTCAACCTCACTAC 859
QY 663 TGCCTTCCAGTCTCTGAAGAGATCCACAGGAGGCACTTGTGCTAGCTGTAAAGT 722
DB 860 TGTCTTCCAGTCTCTGAAGAGATTAATCAAGAGGATTTCAACCTCACTAC 919
QY 723 GGAAGTAAAGATTTTCTTGGAGCTAGATGATGCTCTCAATGATGAGCAAGAAAG 782
DB 920 GGAAGTAAAGATTTCTTGGAGCTAGATGATGCTCTCAATGATGAGCAAGAAAG 979
QY 783 TGTCTTGTGATGATGCTGCTGATTTTCAATGATGATGATGATGATGATGATG 842
DB 980 ATATCTGCTGAGTGTGCTGATTTTCAAGCTCTCAATGATGATGATGATGATG 1039
QY 843 GTGTTTGAAGAAATGCACTGCAAGCTTGAAGATTTCACTAGCTTCAATGACCA 902
DB 1040 GTCTTTGATGAGGCGGCTGCTGCAAGATTTGATGATGATGATGATGATGATG 1099
QY 903 GATTTTATGAGCTTCTTGAAGCACTCAAGATTTAAGTGAAGAGCCATGAGAG 962
DB 1100 GATTTTATGAGCTTCTTGAAGCACTCAAGATTTAAGTGAAGAGCCATGAGAG 1159
QY 963 GTACCCGAATCTTTTCTTATGATGATGATGATGATGATGATGATGATGATG 1022
DB 1160 GTTCCAGAGCTTTTCACTTCCATGATGATGATGATGATGATGATGATGATG 1219
QY 1023 CTCGATGCTGTCGCG 1037
DB 1220 CTTCAATGCAACCG 1234

RESULT 14
CN50A2XU 1297 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLT932B09 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX828087
VERSION 1
KEYWORDS HTG; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1297)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS Castelli V., Aury J.M., Jallion O., Wincker P., Clepet C.,
Menard M., Crnaud C., Quetier F., Scarpelli C., Schachter V.,
Temple G., Caboche M., Weissenbach J., Salanoubat M.
TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1297)
Genoscope.
AUTHORS Direct Submision
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Crnaud C., Schachter V., Weissenbach J., Salanoubat M.
URV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information Center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full1length
<http://www.genoscope.cns.fr/cgi-bin/gsb/gsb?source=Arabidopsis>.

FEATURES

source

1. .1297
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSLTPGHS02G03"
 /issue_type="Hormone Treated Callus"
 /plasmid="pCMVSPORT_6"
 1. .1297
 /gene="At4g22930"

gene

ORIGIN

Query Match 45.2%; Score 574.4; DB 3; Length 1297;
 Best Local Similarity 72.8%; Pred. No. 1,7e-147;
 Matches 754; Conservative 0; Mismatches 281; Indels 1; Gaps 1;

QY 3 GCAAAATGAGCTCTCAATCAACAACCGATGATTGGCATCTTCATCCGGATGTT 62
 DB 172 GTAAGATGAGCTCAACATCACTCAACCTGATGATGGCATCTTCATCCGGATG 231
 QY 63 GATGTTCTTAAGGAGCTGCTCTCAAGTCAATCACTTTGGAGGCAATAGTCATG 122
 DB 232 GATCTTCTTCACTGCTGTTGTTCCCAACAGTCCAGATATTTTAAGAGAGGATTTGATG 291
 QY 123 CCAAAATTTGAAGCTCTCTATCACTACCACTGCTGCTGTAGCATACCGGAGGCGATA 182
 DB 292 CCGAATCTGAAGCTCTCTGACCTCCACCTGACAGCTGCCATTTATTTACCGAAATTTATC 351
 QY 183 TTGAAATCTTACTCTGTTGATGATGATTCACCTCTTATGACCTTATTTGACAGAT 242
 DB 352 ATGAAGCTTGGCATCTGAGAGTGTGATGATCACTTATGACATTTATTTGACTGAC 411
 QY 243 ACAACGAGCTCTATGAAATCAAACTAGCAAGAGAGAGCAGCTGTATTTGGGGTAG 302
 DB 412 AAAAATCTACTGAAGAGATCAAGCTTGCAAGGAAAGTGTGTTATGGGGTAG 471
 QY 303 TTGTACCTCTGCTGTCACGACCAAAATTTCTCAAGATGAGTACTGATCTTTTGGGAG 362
 DB 472 CTGTAACCTGCGGAGCAACAACACTCAAGATGATGTCACGAGATCTCTTTGGA 531
 QY 363 TGTTTACCACTTCAACAAGAAATGTTGAGATTAATGCTCTGCTGTTCAATGAGAG 422
 DB 532 TGCTTACCGGTACTGAAAGATGTTCAAAACAAATGCTTTATCTGTTCAATGGGAG 591
 QY 423 GTTACTATCTCTGAGTTGACATGTTTATGAGAAAGATTCATTGAAACGGTCTA 482
 DB 592 GTCAAGATCCGATATGATGATCTTTGACCGGAGAAATCTTATGAGACAGTCTG 651
 QY 483 AGACCGTGTGTCAGAAATTTCAAAATGAAGGTGATGAGACATTTTACCAACAT 542
 DB 652 CAGCTCTAATCCAAAGCCTTCCGAGCTGAAAGATGATGAAACATCACAAACATG 711
 QY 543 GATGCTGTAAAGTTTGTGGAATCTTGCACTGAAGATTTGTGACGAACTGTACCCCA 602
 DB 712 GATGCTGTAAATTTGTGGAATCTTGCAAAAGAGGTGTGGGTGCAAAAGTACACCA 771
 QY 603 CAACATCTGTTTGAAGAGATTTCTCTTCCAAAGGGGCTTACCAACCCATATATAC 662
 DB 772 CAACATCTCTTCAACAGAGAGAGCCCTTTTCCAAAGGTGATTAACACCTCAACAAC 831
 QY 663 TGCCTTCAGTCTCAAAAGAGATTCACAGGAGGACATTTGTCTCACTGTAAAGAT 722
 DB 832 TGTCTTCCGTTCTCAAAAGAGAAATCATAGAGAGCAATTTGCAAAAGCTGTAACTAGT 891
 QY 723 GGAAGTAAAGATTTTCTTGGGACATGATGCTCTCATGATAGCGAAGAAAGAG 782
 DB 892 GGAAGCAAGAGTTCTTCTCGGAGCAATAGTGTCTCAACATGACGAGTAGAAGAA 951

QY 783 TGTCTTGTGATGATGCTGTGATTTTACAAATGACACCTGTAGCCTTG-TCAGTATAGCGAA 841
 DB 952 TATCTCTGGAGATGCTGTGATTTTACAGAGCTCCATTCCTTCTCTTACGCCAA 1011
 QY 842 GGTGTTGAAAAGAAATGACATCGACAACTTGAAGCATTCATGATTCATAGGACC 901
 DB 1012 GGTCTTTGATGAGGGGGGTGCGCTGACAAAGTTGAAAGCTTTCACAGTTTCAATGACC 1071
 QY 902 AATTTTATGAGGCTTCTTGAAGACATCAAAAGATTAAGTTAGTTAAGACCCATGAA 961
 DB 1072 TATTTCTATGAGCTTCCGAGAAATCTGCAAAAGATCAACGAAAGATTTCTCTTGAA 1131
 QY 962 GGTACCCGAAATCTTTTCTTATGATCAGAGATATATTCATGTTGCTGTGTAAT 1021
 DB 1132 GGTTCAGAGCTTTTCACTTCCATTCGAGAGATGCTCTTATGTTGCTGAGAAAC 1191
 QY 1022 GCTGACCTGTTGCCG 1037
 DB 1192 CTTCAATGCAACCG 1207

RESULT 15
 CK251764 792 bp mRNA linear EST 30-JUL-2004
 LOCUS CK251764
 DEFINITION EST735401 potato callus cDNA library, normalized and full-length
 Solanum tuberosum cDNA clone POCBP96 5' end, mRNA sequence.

ACCESSION CK251764
 VERSION CK251764.1
 KEYWORDS GI:39805104
 EST.

SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Bukarjoca; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 792)
 Buell, C.R., Hart, A., Zismann, V., Karanicheva, S.A. and Baker, B.
 Generation of ESTs from potato callus tissue

TITLE Unpublished (2003)

JOURNAL The Institute for Genomic Research
 Contact: Robin Buell
 9712 Medical Center Dr, Rockville, MD 20850, USA

COMMENT Email: potatc-array@igrr.org
 Clones can be requested from the University of Arizona Genomics
 Institute via <http://genome.arizona.edu/orders/>.
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1. .792
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POCBP96"
 /issue_type="callus"
 /lab_host="DH10B-TonA"
 /clone_id="potato callus cDNA library, normalized and
 full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Solanum tuberosum var.
 Kennebec callus tissue grown on solid media."

ORIGIN

Query Match 44.1%; Score 561; DB 7; Length 792;
 Best Local Similarity 100.0%; Pred. No. 7.5e-144;
 Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGCAAAATGAGCTCTCAATCAACAACCGATGATTGGCATCTTCATCCGTGATG 61
 DB 232 TGCAAAATGAGCTCTCAATCAACAACCGATGATTGGCATCTTCATCCGTGATG 291
 QY 62 TGATGTTCTTAAGGAGCTTGTCTCTCAAGTGCATCATCTTTGGAGGGCAATAGTCAT 121
 DB 292 TGATGTTCTTAAGGAGCTTGTCTCTCAAGTGCATCATCTTTGGAGGGCAATAGTCAT 351

```

QY      122  GCCAATTGAGCCCTCTATGACCTACGCTGCTGCTGATGACATACCGGAGCGAT 181
Db      352  GCCAATTGAGAGCCCTCTATGACCTACGCTGCTGCTGATGACATACCGGAGCGAT 411
QY      182  ATTGAAATCTTTACCTGTTGATGATGATTTCAACCCCTCTTATGACACTTTATTGACAGA 241
Db      412  ATTGAAATCTTTACCTGTTGATGATGATTTCAACCCCTCTTATGACACTTTATTGACAGA 471
QY      242  TACAACCAAGTCTATGAAATCAAACTGCAAGAGAGAGCCAGTCTGATTTGGGGTGA 301
Db      472  TACAACCAAGTCTATGAAATCAAACTGCAAGAGAGAGCCAGTCTGATTTGGGGTGA 531
QY      302  GTTGACCTGCTGCTGCGCAGCAAAATTCCTCAAGATGAGTGA CTGATCTTTCCGGA 361
Db      532  GTTGACCTGCTGCTGCGCAGCAAAATTCCTCAAGATGAGTGA CTGATCTTTCCGGA 591
QY      362  GTGTTACCAAGTTCTACAGAAATGTTGAGCATATATGCTGCTGCTGTTCAATGAGA 421
Db      592  GTGTTACCAAGTTCTACAGAAATGTTGAGCATATATGCTGCTGCTGTTCAATGAGA 651
QY      422  GATTACTAATCCTGAGTTGACATGTTGATAGAGAAAGTATTCATTGAAAGGTTCT 481
Db      652  GATTACTAATCCTGAGTTGACATGTTGATAGAGAAAGTATTCATTGAAAGGTTCT 711
QY      482  AAGACCGTTGGTGAGAAATTTCCAAATTGAAGTGTGATGAGCATGTTACCAACAT 541
Db      712  AAGACCGTTGGTGAGAAATTTCCAAATTGAAGTGTGATGAGCATGTTACCAACAT 771
QY      542  TGATGCTGTAAAGTTTGTGA 562
Db      772  TGATGCTGTAAAGTTTGTGA 792

```

Search completed: February 12, 2005, 05:08:17
 Job time : 4627 secs